DOCUMENT RESUME

ED 057 807 LI 003 296

AUTHOR Forrest, Kathryn S.
TITLE Center for Information Services, Phase II: Detailed System Design and Programming, Part 2 - A Study of Customized Literature Searching Using "CA Condensates" and "CT" Magnetic Tape Data Bases, Phase IIA Final Report.
INSTITUTION California Univ., Los Angeles. Inst. of Library Research.
SPONS AGENCY National Science Foundation, Washington, D.C.
PUB DATE 1 Mar 71
NOTE 89p.;(6 References)

EDRS PRICE MF-$0.65 HC-$3.29
DESCRIPTORS *Computer Programs; *Data Bases; *Information Centers; *Information Retrieval; *Information Services; Search Strategies
IDENTIFIERS *Computer Software; University of California (Riverside)

ABSTRACT

An experimental program using two magnetic tape data bases from Chemical Abstracting Service ("CT," and "CA Condensates") at the University of California at Riverside (UCR) is reported. Cost studies are based entirely upon the University's ability to purchase the tapes and the Computing Center's interest in underwriting part of the project. Sophisticated cost analysis and possible income were not considered to be major objectives of the project; thus, some study information is somewhat superficial and poorly documented. The basic goal was service to the UCR campus clientele. Some of the problems are a result of resistance from within the organization both by users through fear of failure and expense and by the administration in their inability to provide adequate staff and funding. In spite of these problems, a working system for both "CA" and "CT" tape services has been developed, and some experience and expertise in profiling and teaching methods, and a growing number of satisfied users have been gained. (Related documents are LI 003295 and LI 003297 through LI 003301). (Author/NH)

ED057807

U.S. DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
OFFICE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

CENTER FOR INFORMATION SERVICES

PHASE II: DETAILED SYSTEM DESIGN AND PROGRAMMING

NSF GRANT GN-827

PHASE IIA FINAL REPORT

PART 2

A STUDY OF CUSTOMIZED LITERATURE SEARCHING USING CA CONDENSATES AND CT MAGNETIC TAPE DATA BASES

by

Kathryn S. Forrest

1 March 1971

Institute of Library Research
University of California
Los Angeles, California

LI 003 296

TABLE OF CONTENTS

		PAGE
	LIST OF ILLUSTRATIONS.	iii-iv
	AUTHOR PREFACE AND ACKNOWLEDGEMENTS.	v-vi
I.	INTRODUCTION	1
II.	PLANNING AND RATIONALE	3
III.	ORGANIZATION AND ORIGINAL EXPERIMENTATION.	12
IV.	CT AND CA CONDENSATES MAGNETIC TAPE SERVICES . . .	15
V.	PROFILE PREPARATION AND SEARCH PROCEDURES.	17
	Operation and Work Flow	17
	Seminar Development	21
	Description of Data Elements and Search Fields.	25
	Description of Search Strategy.	33
VI.	EVALUATION WORKSHOPS AND INTERVIEWS.	47
VII.	PRINTABLE DATA ELEMENTS AND FORMAT	54
VIII.	COST ESTIMATES AND FUNDING	57
IX.	RESULTS AND CONCLUSIONS.	60
	EPILOGUE	72
	APPENDIX A	78

LIST OF ILLUSTRATIONS

FIGURE		PAGE
1	CHEMICAL ABSTRACTING SERVICE. SEARCH PROGRAMS AND DOCUMENTATION PROVIDED FOR IBM 1401 and S/360 . .	5- 6
2	PANDEX. MARC II FORMAT	7
3	ISI (Institute for Scientific Information) ASCA II (Automatic Subject Citation Alert).	8
4	COMPARATIVE STUDY OF MAGNETIC TAPE SERVICES BASED ON JOURNALS IN THE BIO-AGRICULTURAL LIBRARY UNIVERSITY OF CALIFORNIA, RIVERSIDE	10-11
5	ORGANIZATIONAL CHART - 1968	13
6	GENERAL PROGRAM FLOW CHART.	18
7	LIBRARY AND USER FUNCTIONS.	19
8	COMPUTER CENTER FUNCTIONS	20
9	SEMINAR PUBLICITY SAMPLE.	22
10	COURSE OUTLINE.	23-24
11	SEARCH PROFILE FORM	27
12	CA CONDENSATES CITATION	27
13	CA CONDENSATES CITATION ILLUSTRATING APPLIED TEXT WORDS	27
14	CODEN	29
15	PATENT CODEN AND ABBREVIATIONS OF COUNTRIES	30
16	CHEMICAL ABSTRACTS (ODD) Sections: 1-34	32
17	CHEMICAL ABSTRACTS (EVEN) Sections: 35-80	32

LIST OF ILLUSTRATIONS (Continued)

FIGURE		PAGE
18	PREPARED PROFILE.	35-36
19	SEARCH GUIDE.	37
20	TERM TRUNCATION MODE - EXPLANATION.	42
21	TERM TRUNCATION MODE - EXAMPLE.	42
22	WEIGHTING FACTOR CODING ON SEARCH PROFILE FORM. . . .	45
23	WEIGHTING FACTOR CODING ON PRINTOUT FORM.	46
24	INSTITUTE OF LIBRARY RESEARCH. CENTER FOR INFORMATION SERVICES. SEMINAR. EVALUATION WORKSHOP.	48-50
25	USER FEEDBACK RECORD.	51
26	SDI UPDATE.	51
27	CT PRINTOUT FORMAT (CAS).	51
28	CA CONDENSATES PRINTOUT FORMAT (CAS).	55
29	SAMPLING OF COST ANALYSIS PER PROFILE	58
30	TWO YEAR COMPARATIVE INCOME AND EXPENDITURES.	59
31	KEYWORDS APPLIED FROM ABSTRACT.	64
32	CT PRINTOUT FORMAT - UCR.	67
33	CA CONDENSATES PRINTOUT FORMAT (UCR).	68-69
34	BA PROFILE FORM - UCR	73
35	BA PROFILE PRINTOUT - UCR	74
36	BA PRINTOUT FORMAT - UCR.	75
37	BIORI PRINTOUT FORMAT - UCR	76
38	BILLING RECORD - UCR.	77

AUTHOR PREFACE AND ACKNOWLEDGEMENTS

I ask the indulgence of the reader and hasten to explain that the purpose of this document is to partially satisfy the information needs of three audiences with somewhat divergent goals. First, as indicated by the title sheet, this report was prepared as a formal part of the final Phase IIA progress report to the National Science Foundation outlining our part in GN-827: Development of a Center for Information Services Phase II: Detailed System Design and Programming. Phase IIB of this study continues at UCLA with joint participation by the Institute of Library Research, the UCLA Campus Computing Network, and the UCLA Library. The second intent was to prepare a detailed manual for use by reference librarians and perhaps some individual users within the University of California. The third goal, incidental to the aforementioned purposes, was to document my participation in this project and transmit to my successor at UCR a history and simple treatise for continuation, further investigation and experimentation with the Chemical Abstracting Service magnetic tapes in their new format. What may prove to be adequate knowledge for one audience may be interpreted as important or trivial to another. While the subject is of wider scope than can be covered effectively in this document, there is a need for more detailed information on profile building and preparation. This report is an attempt to meet that need.

I would like to express my heartfelt thanks to Dr. Robert W. Rex, a Geophysicist at UCR, who expressed the original need for SDI service on our campus and encouraged my involvement in the project.

I am indebted to Dr. Morris Garber, Director of the UCR Computing Center for his cooperation and financial support in providing machine time for research and development.

I owe very special thanks to the Chemical Abstracting Service for their manuals from which I so freely quoted and without whose magnetic tape service the project would have been impossible.

It is with deep appreciation that I acknowledge the untiring efforts of my secretary, Diane Bennett, who not only gave her time to typing the manuscript, but who has also been involved from the outset with the clerical and statistical activities of the project in addition to her other duties.

Lastly I wish to express my gratitude to Mr. Robert L. Carmichael, Manager of the Institute of Library Research, University of California, Los Angeles, who edited the final manuscript following my resignation from the University of California.

Kathryn S. Forrest
Assistant University Librarian, UCR
Retrieval Project Leader
February 26, 1971

I. INTRODUCTION

During early 1968 the Riverside campus became greatly concerned with the proliferation of recorded scientific knowledge and with the researchers' inability to quickly extract the relevant data. Although widely discussed in the literature during the previous five years, very little had actually been done toward document retrieval except in the smaller, more affluent industrial libraries where technical report systems were under development. The Bio-Agricultural Branch Library on the Riverside campus of the University had been using a KWIC (Keyword-in-Context) indexing system since 1964 for reprint and report files; therefore, automation was a natural outgrowth of earlier developments.

We were quite impressed by Saul Gorn's description of the impact of the computer age upon education and teaching. He stated so accurately that "Not only are we forced to face an educational policy question as to what we should or should not permit to be mechanized in education, but also even broader questions of how to approach flexibility in education to prevent technical obsolescence of manpower, and therefore how to counteract the seeming demands for specialization in a complex civilization"(1). Roger Smith went on to say that "there are no borders or 'fences' around specialization. All of them reach into or overlap in part with other fields . . . we build upon the work of others. No teacher or research worker can be most useful or most effective unless he makes a constant effort to keep up to date in his field"(2). According to the Encyclopedia Britannica we must now teach our students 10% more raw material than existed just one generation ago.

Academic concern at Riverside became evident when the Librarian was approached by a faculty group in conjunction with the Director of the Computing Center at UCR. General concensus implied 'the time is now and the library must be the focal point of this new literature technology'.

---

[1]Gorn, S. "The Computer and Information Science and the Community of Disciplines", Behavioral Science, V. 12, 1967, p. 433-452.

[2]Smith, R. C. Guide to the Literature of the Zoological Sciences, 7th ed., Minneapolis, Burgess, 1967.

Our original project outline developed as a long-term, five phase approach primarily based on anticipated purchasable data bases in the sciences, but also heavily relying upon the use of MARC II input and in-house magnetic tape data base preparation. Our original objectives were for in-depth indexing of all relevant scientific knowledge within our UC campus or readily available on Interlibrary Loan. By mechanical means we proposed to retrieve source data at regular intervals for our faculty and students to reduce the heavy burden of search time through indexes, books, and journals.

This approach has been regularly refined as knowledge expanded and data bases quickly developed, paradoxically, in parallel with budget cuts and staff freezes. This report then does not pretend to describe the level of operation or success which we are confident could have been accomplished with even nominal budgetary support. This report does explain how one institution experimented with magnetic tape data bases, maintained a respectable level of contributive research and development, taught profile preparation, and ran an operational SDI system for multi-campus clientele with small additional staff and funding.

The primary objective of this report, however, it to discuss the procedures, pitfalls and results gained from our experiences with two magnetic tape data bases contracted from Chemical Abstracting Service (CT, and CA Condensates) so that the continued design and development of UC information services can benefit from this added insight.

II. PLANNING AND RATIONALE

This project was and is basically concerned with user-oriented retrieval of information in its broadest context. All research workers know the frustrations of attempting to keep current when the escalation of publication has long since surpassed their reading and retention capacity. Even when a small percent of the basic literature is examined, endless hours of indexing and cross-referencing of reprints are spent in an attempt to improve recall. It has been said that the human brain is a poorly organized storage device and even less useful for rapid retrieval.

Modern methods of hand indexing and individual cataloging will no longer appease the hungry scientist in his insatiable quest for answers. Knowledge is no longer limited to geographic boundaries. Scientific research is no longer the prerogative of three or four major powers and neatly wrapped in two scientific languages. The inevitable time lags between primary publication and conventional abstracting and indexing methods are no longer acceptable.

Armed with this knowledge of our users "need to know" and the enthusiastic and sympathetic support of the campus administration and faculty, investigations began pursuant to available machine-readable magnetic data base services.

For this project, we proposed the use of purchased magnetic tape services covering scientific serial literature as a data base, even though our initial goal was to encompass the needs in all disciplines. Basically the sciences lent themselves to a more current and controlled study since scientific abstracting and indexing firms had long influenced the proliferation of serial literature through subject indexing. Even here the extensiveness of the citations to be published had also caused time-lag gaps, diminishing their usefulness, as these firms also become bogged down in paper. The National Federation of Science Abstracting and Indexing Services released statistics on abstracting and citation coverage provided by its fourteen current member services which showed a 157% increase in the decade 1957-1967. Chemical Abstracting Service had just estimated that 261,000 abstracts in 80 subject classifications would be published in 1969. Bioscience Information Service would reference more than 220,000 research papers during that year. The

Institute for Scientific Information expected to supply data during 1969 at the rate of 5,400 current items weekly. These were the only three services employing subject specialists who were then prepared to supply magnetic tape service.

In April, 1968, I was appointed Project Leader for the actual Retrieval Project at UCR and spent the next several months in research and analysis. Our project developed as follows. Some cost studies were undertaken and budget allocations were discussed. In May, 1968, comparisons were made on cost/coverage of three services which were reportedly available for subscription. These were Chemical Abstracting Service (Figure 1), Pandex (Figure 2), and Institute of Scientific Information (Figure 3). Sample tapes were requested from the first two companies along with whatever programs and documentation they were prepared to supply. The UCR Computing Center agreed to support limited programmer and machine time for experimentation.

A study of duplication of publication coverage among seven advertised data bases was undertaken in respect to our scientific serial holdings at UCR. This was accomplished in July 1968 by checking the service-supplied list of periodicals covered against the UCR library list of serials holdings for the following tape services: Chemical Abstracts (CA Condensates), Chemical-Biological Activities (CBAC), and Chemical Titles (CT), all services of Chemical Abstracting Service; Pandex, Engineering Index, Biological Abstracts, and the Institute of Scientific Information's Automatic Subject Citation Alert (ASCA II).

FIGURE 1

CHEMICAL ABSTRACTS SERVICE
Search Programs and Documentation Provided
For IBM 1401 and S/360

		Cost per Year
CBAC	(Chemical-Biological Activities)	$1,600.00
	600 journals of which we now take 237.	
	Includes printed issues and indexes.	
	Includes in-vitro reactions of organic compounds.	
	Began with 1964. (21,000 titles plus abstracts).	
	Requires 7 tapes to hold data to 1968.	134.75
	Subscribers provide magnetic tapes.	
	9-track 2,400 ft. tapes recorded at 800 bpi.	
	Blank tape $19.25 each in lots of 20 (through UCR Computing Center).	
	A tape is prepared every 2 weeks.	
	We do not subscribe.	
CHEMICAL TITLES		1,500.00
	Generates KWIC type index.	
	650 journals of which we now take 286.	
	Includes printed subscription.	
	File dates from Jan. 1962. (500,000 titles in 1967).	
	Requires 5 tapes (as above) for data in 1968.	96.25
	A tape is prepared every two weeks.	
	Our subscription cost. (Hard copy free with tape service).	-50.00
POST	(Polymer Science and Technology)	
	POST J (Journals and Reports)	1,700.00
	POST P (Patents)	1,500.00
	POST J & P (Both)	3,100.00
	Tapes to POST J and POST P are prepared on alternate weeks.	
	Coverage: Research, development, production, equipment and other aspects of polymer chemistry.	
	Tapes contain digests published in printed issues.	
	We do not subscribe.	
CA CONDENSATES	(Both)	5,000.00
	Entire content of CA in machine readable form.	
	Includes: titles, authors, journal, and patent references, and CA issues. Keyword Index for each abstract published.	

FIGURE 1 (Continued)

CA BASIC JOURNAL ABSTRACTS (Each) 4,000.00

Abstracts are selected from journals covering broad range of chemical interest.
15,000 full abstracts will be included on tape in 1968.
Covers 34 journals.
Began March 1, 1968.

K. Forrest
May 1968

FIGURE 2

PANDEX
MARC II Format

Coverage: Science, Technology and Medicine.
1. 2,000 key periodicals (all articles).
2. Begins calendar year 1967.

Format:
1. Quarterly index.
2. Cumulated annual is 4th quarterly.
3. 4 x 6 microfiche.
   (a) 27 columns of index entries.
   (b) Each column has "header" term.
4. Index issued in 2 parts.
   (a) Subject
   (b) Author (includes primary and all secondary authors).
      (1) All bibliography information given with primary author.
      (2) Journal, paging and cross reference to primary author given with secondary authors.
      (3) Periodical titles are given in ASTM coden form.
      (4) A sheet of full titles and coden is given with index.
      (5) Subject index is KWIC-type and human indexer subject heading combination.

SDI Capabilities: Available November 15, 1967
1. Magnetic tapes generated on IBM 360 either 7 or 9 track format.
2. Available after November 15, 1967, on weekly basis.
3. Annual subscription rate--$6,500--includes search program and full documentation.

Future:
1. Plans underway to index all English language books published in science, technology and medicine (approximately 6,000 titles a year). This is expected to begin January 1, 1968.
2. Plans underway to include Engineering Index (Electronics and Plastics Section) beginning July 1, 1968.

FIGURE 3

ISI (INSTITUTE FOR SCIENTIFIC INFORMATION)
ASCA II (Automatic Subject Citation Alert)

1. Tapes delivered weekly
   a. Approximately 5,400 current items each week.
   b. Time lag between primary publication and tape delivery is 1-3 weeks.
   c. Comprehensive coverage of all items except ads and ephemeral notices.
      (1) 1,600 key scientific journals in 1967 of which we take 816.
      (2) 689 are duplicated by coverage in Pandex or Chemical Titles.
   d. Coverage available from 1964.
   e. Includes all current U. S. patents issued (primary and reissue).

2. Tape description
   a. Compatible with IBM 360.
   b. Ordinarily produced in 556 bpi density.
   c. Available in 7 or 9 channel.

3. Cost
   a. Source tapes.
      (1) $5,000 to $8,000 depending on journals selected.
      (2) $10 per user over first 100 users for selective dissemination purposes.
      (3) Variable length blocked records.
   b. Citation tapes.
      (1) $15,000 to $24,000 per year depending upon source journals selected.
      (2) Same as above.
      (3) Base price adjusted according to user's own statement of fraction of source journals scanned by user.
      (4) Fixed-length records.

These sources were checked and known Coden was annotated: From the very simple method employed (Figure 4) we determined the total journal coverage within our campus library system and the subject disciplinary structure. We could also determine to some extent the University-wide and external impact upon Interlibrary Loan operations. For example, we determined that UCR subscribed regularly to 43.2% of the total titles in CT. Even if our chemists could read each of these journals on a regular basis, they would not be currently aware of other research covered in 56.8% of the basic world literature. UCR subscribed to 6,774 current serial titles in June, 1968, a 12.45% increase over June, 1967, and the combined University of California libraries subscribed to 148,508 serial titles, a total increase of 10.60% during the same period. It was surprising to find that even with this large staggering number of combined serial holdings, the Bio-Agricultural Library at UCR alone requested 14.12% of their interlibrary loans from libraries outside the University System.

Findings showed that Chemical Abstracting Service coverage would best meet the needs of our primary investigation. This was determined by the knowledge that our widest journal coverage was included in two of three services and than an unusually large group of chemists were employed in virtually every scientific department on the campus. We selected CT, the magnetic tape service name for the hard copy edition known as Chemical Titles, and CA Condensates, the magnetic tape service name for Chemical Abstracts. With these two services from Chemical Abstracting Services, we felt that our scientists and students would be adequately informed of world-wide research and development in the multi-faceted field of chemistry.

Based on these findings and the implications of related problems besetting the entire academic community, a decision was made to continue with an experimental program using these two magnetic tapes. It must be quickly pointed out that at this time our cost studies were based entirely upon our ability to purchase the tapes and the Computing Center's interest in underwriting part of the project since we were all tremendously interested in the anticipated impact of information retrieval capabilities within library operations. Sophisticated cost analysis and possible income were never considered to be major objectives of our project proposal; consequently, some study information is somewhat superficial and poorly documented. It must be stated here that the basic goal at this stage was service to our own campus clientele at UCR.

FIGURE 4

COMPARATIVE STUDY OF MAGNETIC TAPE SERVICES BASED ON
JOURNALS IN THE BIO-AGRICULTURAL LIBRARY
UNIVERSITY OF CALIFORNIA, RIVERSIDE

		Alphabetical List of Titles	Coden	Chemical Abstracts	Chemical-Biological Activities	Chemical Titles	Institute for Scientific Information	Pandex	Engineering Index	Biological Abstracts
1		A.I.A.A. Journal	AIAJ					X	X	
2	E	A.I. CH.E. Journal	AICE	X		X		X	X	
3	G	Academie Des Sciences, Paris. Comptes Rendus Hebdomadaires des Seances. Series A et B.	CHDA	X	X	X		X	X	
4	C	Academie Des Sciences, Paris. Comptes Rendus Hebdomadaires Des Seances. Series C. Sciences Chimiques.	CHDC	X	X	X	X	X	X	
5	BA	Academie Des Sciences, Paris. Comptes Rendus Hebdomadaires Des Seances. Series D. Sciences Naturelles	CHDD	X	X	X	X	X	X	
6	BA	Academy of Natural Sciences of Philadelphia. Proceedings.	PANP	X			X	X		X
7		Acta Antomica.	ACAT	X			X	X		X
8	M	Acta Anatomica. Supplement.		X				X		X
9	C	Acta Biochimica Polonica.	ABPL	X	X	X	X	X		X
10	BA	Acta Biologica.	ABAS		X	X		X		X
11	BA	Acta Biologica. Supplementum.	ABAS					X		X
12	BA	Acta Botanica Neerlandica.	ABNRAN	X	X		X			X
13	C	Acta Chemica Scandinavica.	ACSAA4	X	X	X	X	X		X

FIGURE 4 (Continued)

		Alphabetical List of Titles	Coden	Chemical Abstracts	Chemical-Biological Activities	Chemical Titles	Institute for Scientific Information	Pandex	Engineering Index	Biological Abstracts
14	PS	Acta Crystallographica.	ACCRA9	X		X	X	X		
15	PS	Acta Crystallographica. Supple.	ACYS					X		
16	M	Acta Genetica et Statistica Medica.	AGSM	X			X	X		X
17	M	Acta Genetica et Statistica Medica. Supple.	AGSU	X				X		X
18	C	Acta Histochemica.	AHISA9	X	X	X	X	X		X
19	PS	Acta Mathematica.	ACMT					X		
20	E	Acta Metallurgica.	AMETAR	X		X	X	X		
21	BA	Acta Microbiologica.	AMAHA5		X			X		X
22	BA	Acta Pathologica et Microbiologica Scandinavica.	APMIAL	X	X	X	X	X		X
23	BA	Acta Pathologica et Microbiologica Scandinavica. Supplementum.	APMUAN	X	X	X		X		X
24	PS	Acta Physica Austriaca.	APAS	X			X	X		
25	PS	Acta Physica (Hungarica).	APAHAQ	X		X	X	X		
26	M	Acta Physiologica Latinoamericana.	APLTAF	X	X		X	X		X
27	M	Acta Physilogica Polonica.	APYPAY	X	X					X
28	M	Acta Physiologica Scandinavica.	APSCAX	X	X	X	X	X		X
29	M	Acta Physiologica Scandinavica. Supple.	APSSAD	X	X	X		X		X
30	M	Acta Poloniae Pharmaceutica.	APPHAX	X	X	X				X
31	PS	Acta Polytechnica Scandinavica. Math. and Comp. Machin. Series.	APSM	X				X		

III. ORGANIZATION AND ORIGINAL EXPERIMENTATION

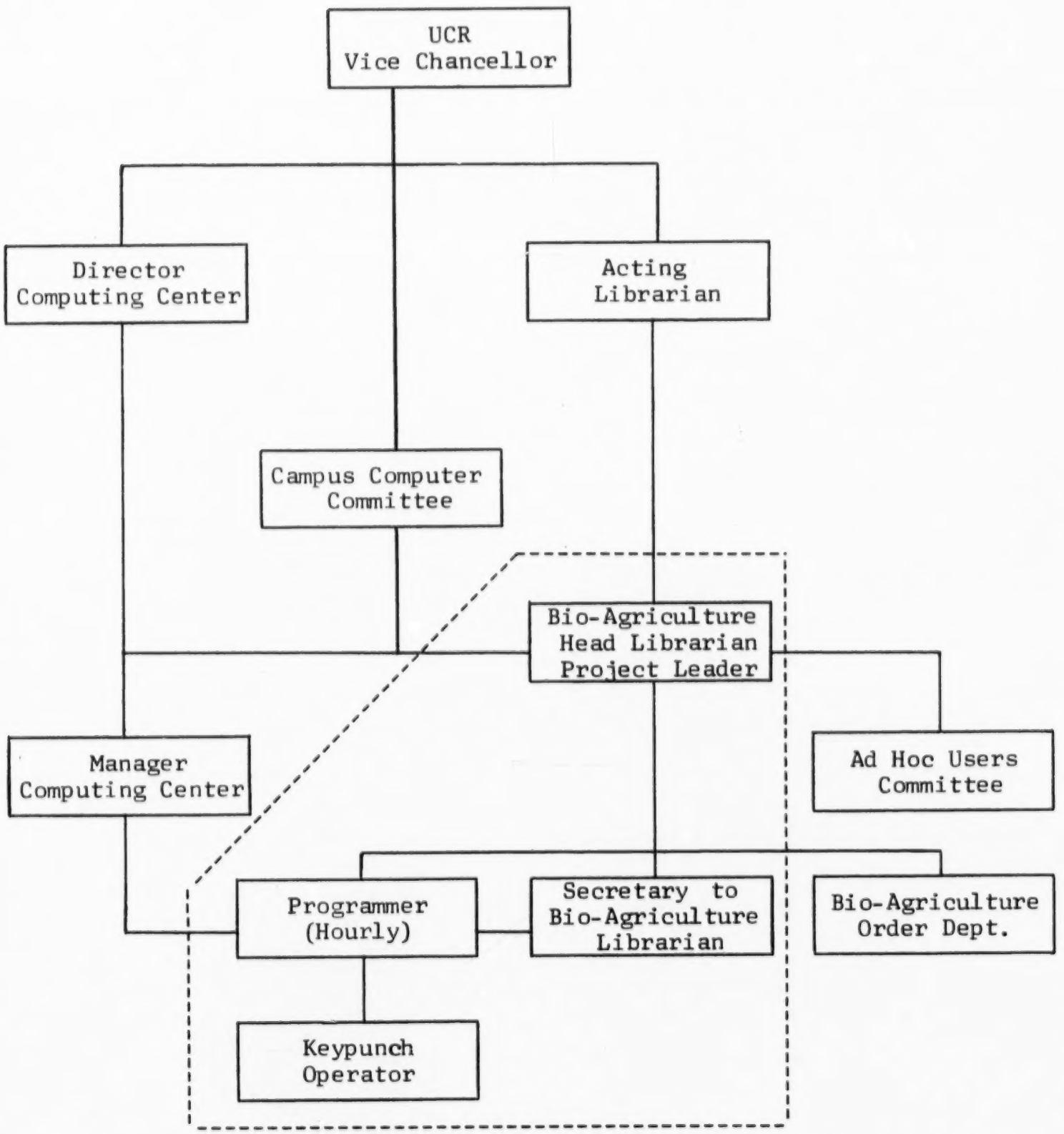
This project was organized within existing operations with responsibilities added to existing positions where needed. Figure 5 shows the original organizational chart of 1968 with the broken lines indicating the major area of activity, which was centered in the Bio-Agricultural Library. The Ad Hoc Users Committee appointed by the Bio-Agricultural Librarian included faculty representing the seven scientific departments, the Computing Center, and both science branch libraries. This committee assisted the Project Leader in determining and resolving issues such as: usefulness of tape services for subject needs; scientist vs. librarian viewpoints for print format methods and descriptive requirements; feedback on system design and cost analysis; keyword application and thesaurus building and/or selection for their various disciplines. The committee was most enthusiastic.

Subscriptions were placed to begin in January, 1969. Arrangements were made with the Campus Computing Center to provide a programmer and keypunch operators, and to recharge only those hours actually used; thus it was not necessary at that time to augment the staff within the library. It is most important to note that only these part time hours were added to the library budget expense for staff. The only other cost to the library was for data base subscriptions and minor supply purchases. Tape acquisitions and payment were carried out through the Order Department with tape expenses being partially paid from a staff publication account augmented from the Bio-Agricultural Library's book fund account.

Our familiarity with the subject matter and hard copy content made analysis somewhat simpler. The software package originally provided by Chemical Abstracting Service with their tapes was adopted. While the programmer analyzed and experimented with the computer oriented system the Project Leader studied the profile preparation materials provided by CAS and compared computer printout results with the hard copy data using personally prepared profiles. We also conferred with representatives of CAS and eventually encouraged the Ad Hoc User Committee to prepare profiles of their own interest for free experimental printouts. These experimental profiles began with the January 1969 issues of CA and CT and continued to run for seven months with regular additions and changes teaching us all new and intricate profile procedures. The secretary

FIGURE 5

ORGANIZATIONAL CHART - 1968



quickly learned to examine newly submitted profiles for technical inaccuracies. She also accepted complete responsibility for collecting selected statistics and mailing out all profile printouts. She also kept comparative records of hard copy versus magnetic tape receipts.

Other campuses of the University became interested in our SDI project and explanatory seminars were conducted on two campuses at their request. Contract negotiations between the University representatives and CAS gave legal permission for UCR to offer service to other campuses of the University. By July 1969, our primary phase of SDI was operational, and three other campuses were represented by regular profile runs. However, research and development continued with the part-time programmer revising and refining the original package to more closely fit user needs.

During the early summer of 1969, discussions began between the Project Leader at UCR and the Director of the Institute of Library Research at UCLA regarding some funding support for both the operational and research and developmental phases of this project. Modest support was provided as part of Phase IIA of the ILR study, Development of a Center for Information Services, Phase II: Detailed System Design and Programming, currently funded by the National Science Foundation on the UCLA campus.

IV. CT AND CA CONDENSATES SERVICE

Subscriptions to the CT and CA Condensates magnetic tape service were requested on 800 bpi, 9 track magnetic tape for use on an IBM 360 computer operating under OS. The tape records were blocked and of variable length. The computer programs were written in Basic OS Assembler language. One IBM 2400 Series magnetic tape drive and three IBM 2311 disc drives were used and the profiles were run by batch processing in a time-sharing environment. An IBM 1403 printer, and IBM 029 keypunch and a verifier completed the necessary equipment.

CT magnetic tapes were supposedly received bi-weekly corresponding to the hard copy issues of Chemical Titles. They were advertised to arrive approximately 35 days prior to the hard copy. This service is primarily intended as a current awareness tool and covers more than 650 basic chemical journals. CT indexes only by KWIC (Keyword-in-Context) indexing methods intended to provide subject analysis from authors' titles, authors' names, and journal coden. Full bibliographic information is included in the citation. We felt that these tapes would have short term value and will probably not be retained indefinitely since each citation should appear in CA Condensates after a lag of three to six months.

CA Condensates appeared weekly corresponding to the hard copy Chemical Abstracts issues. These tapes carry complete bibliographic information including the abstract number. Although the abstract itself is not printed, keywords or phrases from the abstract as well as the title are picked up during a literature search. These tapes were also advertised to arrive approximately thirty days prior to the hard copy.

CA Condensates represented world-wide coverage of chemical literature including some 12,000 journals, conference proceedings, research reports and chemical patents.

Chemical Abstracting Services provided, in addition to a search program with documentation relating to the file, a set of three manuals devised to guide the user through profile preparation. This set consisted of a Preparation of Search Profiles, a Guide to the Effective Use of CAS Search Services on Magnetic Tapes, a Search Guide, For Use with CT, SDI, and CA Issues, and, Chemical Titles Word Frequency List. Chemical Abstracting Service also provided a catalog of Information in December of 1969 which updated the services studied in Figure 1.

Many of the materials provided by CAS in the above named publications were heavily relied on for preparing and teaching profile preparation. In addition to adapting the original program provided by CAS, we also were permitted to duplicate the Search Profile Form, Search System Terminology, and several examples which were combined into a "handout" for use with seminar or individual profile instruction. The set of manuals was also provided through the Reserve Book Desk in both of the UCR Science Libraries. A copy of the "handout" appears as Appendix A.

V. PROFILE PREPARATION & SEARCH PROCEDURES

OPERATION AND WORKFLOW

Chemical Abstracting Service rightfully suggests that "Ideally, the search profiles should be formulated by a person experienced in profile preparation and knowledgeable in the subject matter being searched". At the outset we found it far simpler to attempt teaching profile preparation methods to subject-knowledgeable users than to attempt hiring subject specialists as consultants. This was due to an obvious lack of funds and of personnel experienced in profile preparation. Our first attempt at teaching was directed at highly trained teaching faculty on the Riverside Campus who comprised the Ad Hoc Committee. Through their efforts departmental seminars were given to other faculty and interested graduate students at UCR.

The elementary flow chart on Figure 6 illustrates the primary division of activity within a cooperative library and computing center arrangement. Detailed actions are shown on Figures 7 and 8 for the Library and Computing Center respectively. Step I on Figure 6 (and Figure 7) is concerned with the preparation of the profile through library conducted seminars and revision by the local reference librarian until the edited profile is correct for a free trial run. In Step II (Figures 6 and 8) the Computing Center is responsible for keypunching and editing profiles and accounting data records. These activities are supervised by the part-time programmer who provides a final deck edit. The card deck is then filed to be recalled when the next magnetic tape data bases are received. Step III, also a computing center operation, is initiated after the Library Administration Office has received, recorded and delivered the newly received data base tapes. The part-time programmer prepares the proper input corresponding to the appropriate data base and establishes communication with the computer operator who batches the computer search during non-prime overnight time. The Output from the search operation, which is used in Steps IV through VI, is delivered to the Library Administration Office early the following working day. There clerical functions are completed and profiles are generally in the campus mail to the users (for Step V) by the mid-afternoon pickup.

FIGURE 6

GENERAL PROGRAM FLOW CHART

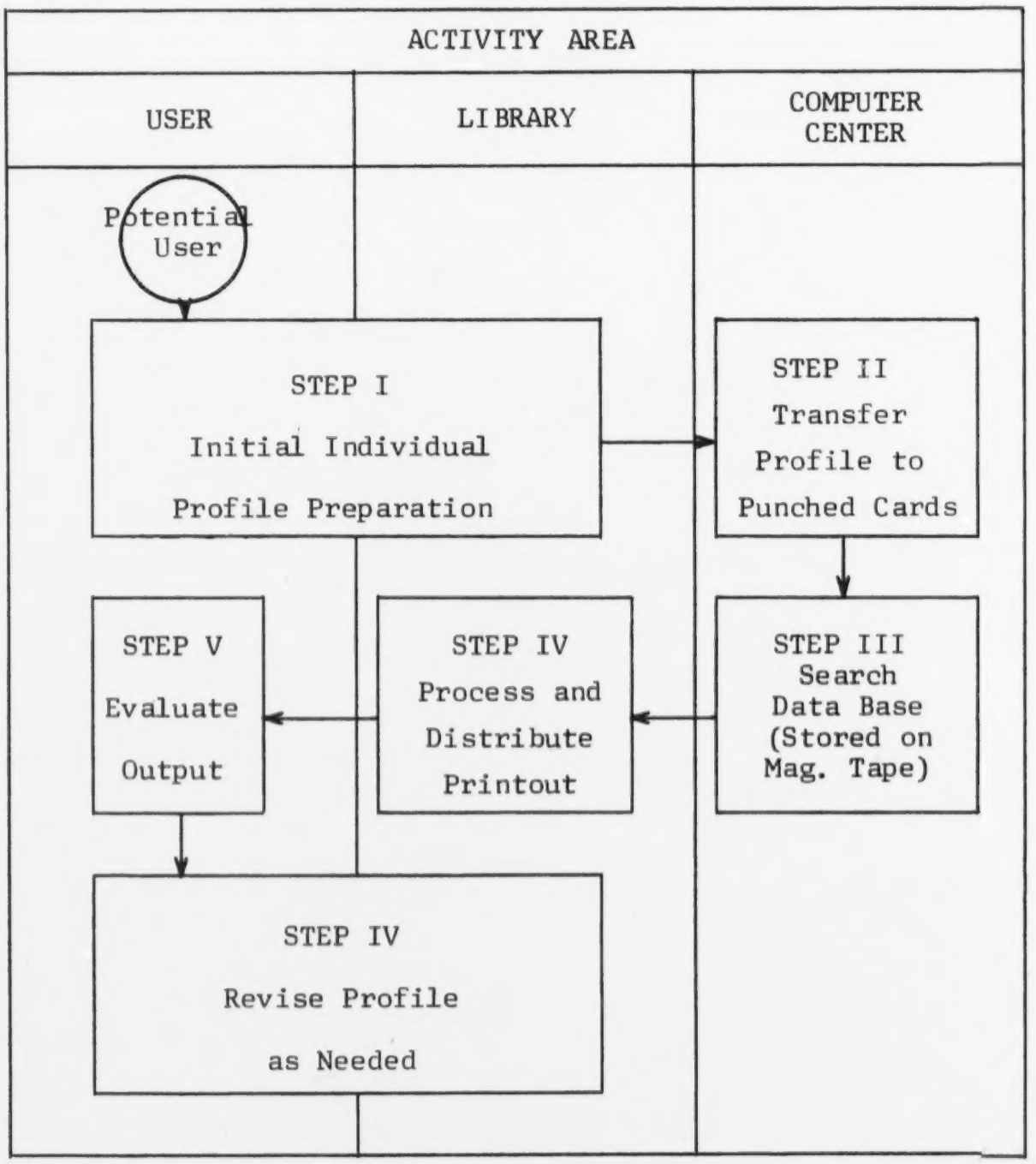


FIGURE 7

LIBRARY AND USER FUNCTIONS

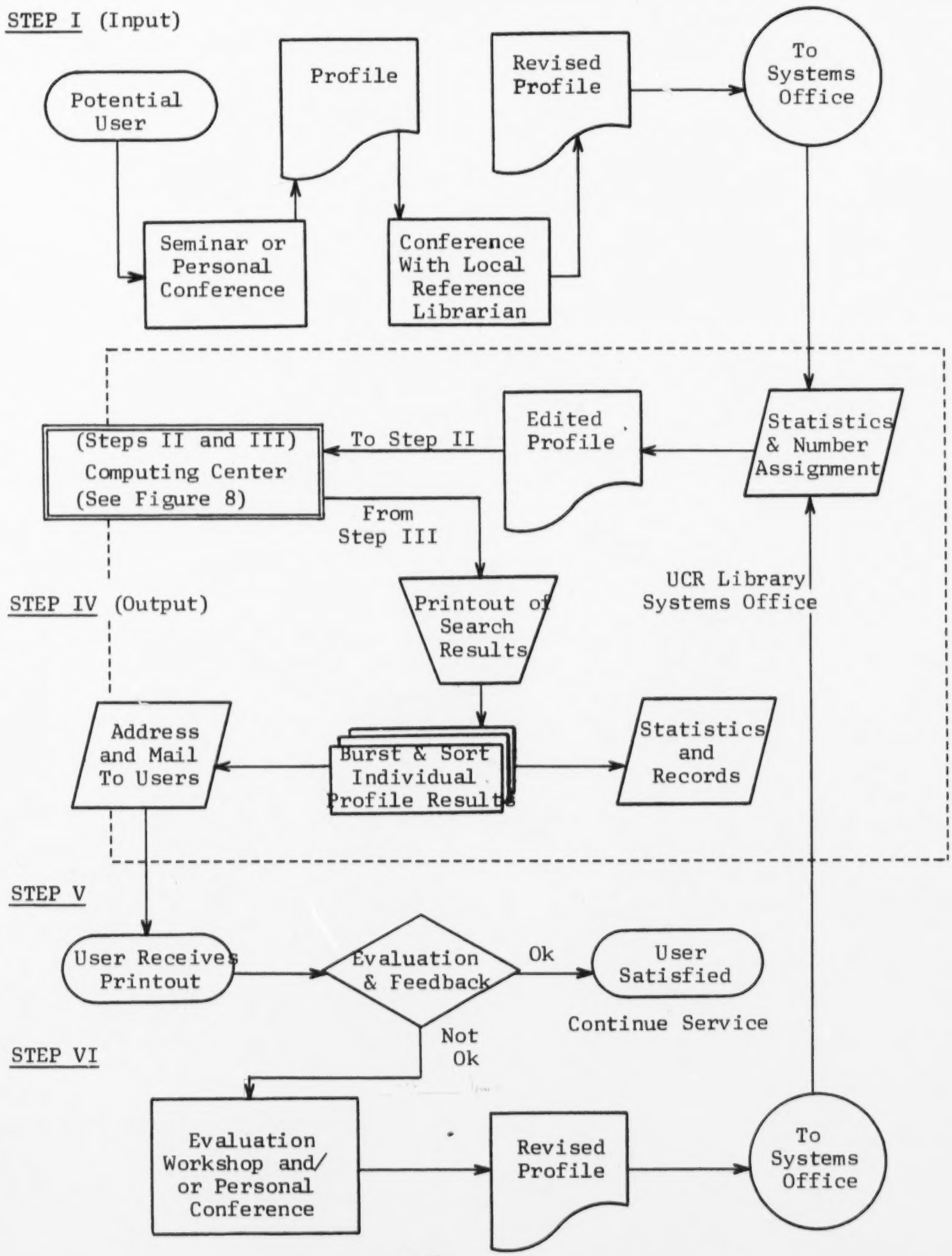
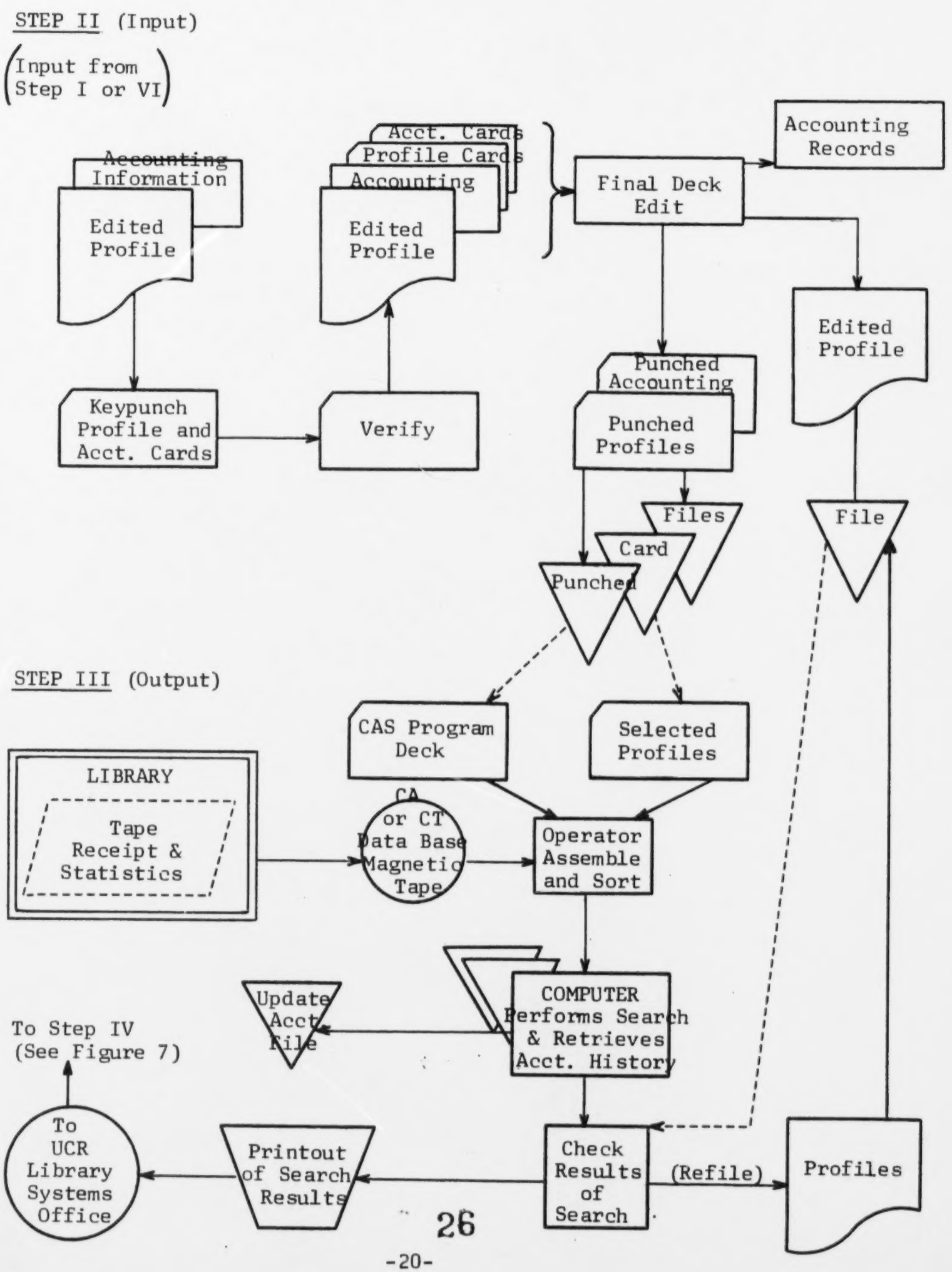


FIGURE 8

COMPUTER CENTER FUNCTIONS



SEMINAR DEVELOPMENT

As observed in Figure 7, the profile preparation seminar is the first step in training a potential user. The basic approach is the same whether teaching a large group, a small seminar or private instruction. Variation occurs only in subject matter and the use of multi-media equipment.

The first few small local seminars helped develop a general outline and quickly pointed up the need for visual explanation. A series of basic transparencies were made to describe the various codes and search logic. Each seminar brought about additions and changes, and at present we have examples drawn from various subject areas. These enhance the usefulness of the seminar to a specialized group of users through the ability to quickly show relevance to their problems.

Local seminars were often presented as a lecture period of a graduate seminar course or sometimes as a departmental seminar including both faculty and students. These were normally advertised through departmental channels and occasionally in the campus calendar. May 1969 brought the first request for a seminar to be held at another campus. That was quickly followed by other requests so that as of this writing eight of the nine campuses have been visited. Seminars offered on other campuses were usually advertised by the local campus science or reference librarian and distributed to each science faculty member. Seminar publicity (Figure 9) varied somewhat with each campus as did attendance and diversity of participants.

It was not until the Institute of Library Research began discussions regarding a general seminar in library computer services to be held at UCLA that a formal seminar outline for the UCR activity was developed (Figure 10). This outline was used only as a guideline and both emphasis and content varied greatly depending upon the audience. (Two of these topics, data elements and search strategy, are discussed in greater detail later in this section).

FIGURE 9

SEMINAR PUBLICITY SAMPLE

M E D I C A L C E N T E R L I B R A R Y

Seminar

USE OF LITERATURE SEARCH SERVICE

Mrs. Kathryn Forrest,
Biomedical Librarian
University of California,
Riverside

3:30 p.m.
Wednesday, May 7, 1969
Toland Hall

A computer-based current awareness service which will cover Chemical Titles, Chemical Abstracts, and Biological Abstracts is now available to Medical Center staff at a nominal charge.*

Description of this service and assistance in its use will be available at the seminar.

Please call a Reference Librarian, x2334 for further information.

(*) Provided on a recharge basis to Medical Center Departments. No other funding is available at this time.

FIGURE 10
COURSE OUTLINE

INSTITUTE OF LIBRARY RESEARCH
Center for Information Services

Seminar on Mechanized Information Service in the University Library

Chemical Abstracting Service Course Outline

I. National and Professional Context

A. CAS as Publisher
   1. Responsibility and Reliability
   2. Hard Copy Preparation
   3. User Groups
B. Terminology and Definitions

II. Input and Subject Authority

A. Tape Services Offered by CAS: Differences, Costs, and Coverages
B. CA & CT in particular
   1. Current Awareness Versus Bibliographic Control
   2. Retrospective Possibilities
   3. File Structure and Type Application
C. Subject Authority Guide
   1. Generic and Grammatical Variations
   2. Term Truncation
   3. Compound Fragmentation
   4. Coden
   5. Word Frequency and Search Guide Lists

III. Data Elements

A. Search Fields
B. Logic and Parameters
C. Format of Output
D. Problems with Weighting

FIGURE 10 (Continued)

IV. Development of Search Strategy

A. Format of Input
B. Coding for Fields and Service
C. Querying User for Keyword Choice and Form
D. Final Revision
E. Keypunching and Statistics

V. Printout

A. Format
B. Data Identification
C. Uses of Data and Prior Verification
D. Costs and Accounting Procedures

VI. Discussion

I want here to add only a few more explanatory remarks regarding the CIS Seminars as they came to be called. These Seminars were sponsored by the Center for Information Services, project of the Institute of Library Research. The first CIS Seminar was held for 20 sessions (two per week) from April through June 1970 and participation was limited to UCLA Library staff members. The second seminar was offered from October through December 1970, also at UCLA, and participants from all five southern campuses attended. Two sessions in each seminar were devoted to the UCR experience with CAS.

As part of our approach to presenting an introduction to the CAS services, we made it possible for each participant to write several profiles for which a printout was received at the beginning of the second class session when evaluation of information and profile restudy was discussed. The impact and results of these seminars are discussed in detail as a separate part of the total CIS Phase IIA final report.

Returning to the discussion of instruction, we found it helpful to put some intrinsic information into the hands of each attendee. This was done by preparing a five page handout primarily copied directly from the Chemical Abstracting Services publication Preparation of Search Profiles. This handout (Appendix A) included the Seven Steps for Coding Search Profiles, Search System Terminology, Mode Symbolic Code, Type Code and a partially completed Search Profile Form which included additional coding information. Blank profile forms and a page of printout showing bibliographic citations and the printed format completed the give-away package. At least one set of Chemical Abstracting Services manuals was deposited with a library on each campus. As previously stated, the transparencies and lectures were regularly shuffled to more closely relate to the subject interest and computer knowledge level of the individual group. This often proved quite difficult, but not impossible.

DESCRIPTION OF DATA ELEMENTS AND SEARCH FIELDS

Much of the teaching became routine assuming a thorough basic background in the hard copy use of both Chemical Abstracts and Chemical Titles and their several indexes. Comprehension of the fundamental reliability of the publisher, the hard copy index configuration and the responsibility gap between publisher and author for keyword continuity and consistency is extremely important.

Search Profile

The Search Profile is a list of keywords provided by the user to describe his search interest. The profile keywords and the subsequent coding provides the specifications on which that specific information is retrieved. The results of a search profile are only as effective as

the quality of preparation of that profile, and an effective long term profile will take considerable initial thought and evaluation.

Figure 11 illustrates the Search Profile Form which is completed by the user and subsequently sent to the keypunch operator. Since the question number if always added at UCR, columns 1 to 3 should be left blank. The output indicators, always 1,000, should be filled in by the user in columns 8 to 11.

Search Fields

Search Fields are areas coded to indicate the line which will be mechanically searched on the magnetic tape. Figures 12 and 13 show sample printout citations from CA Condensates. For the example in Figure 12, the first line designates the tape service followed by the Digest No. (which is the hard copy abstract number) followed by the Chemical Abstracts hard copy volume and issue number. This second line, then, is the bibliographic citation to the secondary source. The reference weight refers to the score of this citation to the search question if a weighting has been applied. Indented under this are the search fields; those areas which can be searched for specific information. The first indented line (1) is the journal coden followed by the volume, issue number, year and beginning and ending page numbers. This is the bibliographic information for the primary publication. The next line (2) is the short title for the journal referred to in the coden line. (Coden will be discussed later as a separate field). Line (3) is the title of the article. The equal sign symbolizes the end of that title. All authors are listed in line (4) and each author's name is searchable. The last line (5) indicates the business address of the first author.

These five search fields are referred to by "type" code symbols placed in column 9 of the Search Profile Form (Figure 11). The following symbolic type codes are used for keywords falling in the associated categories.

Type	Coden
Author's Name	A
Journal Citation	C (Coden)
Patent Citation	C (Special Coden)
Title or Text Word	T

It is necessary to code a key word so as to avoid errors in machine manipulation. For example, the word GREEN could be searched in A as an author's name or in T as a title or text word in searching a color such as GREEN ALGAE.

FIGURE 11

SEARCH PROFILE FORM

Return to:
DIANE BENNETT
GENERAL LIBRARY, ROOM 204
UNIVERSITY OF CALIFORNIA
RIVERSIDE, CALIFORNIA

REQUESTOR: _____

DEPARTMENT: _____

CAMPUS: _____

FUND: _____

Indicate Service
CT____ CA____ BOTH ____
CA-ODD ____ CA-EVEN ____

HEADER CARD

QUESTION NUMBER		OUTPUT INDICATORS	QUESTION WEIGHT
1 3		8 11	12 17
		1 0 0 0	

QUESTION:

DETAIL CARD

DUPLICATE	PARAM NO.	NO. OF PARAM	MODE	TYPE	LOGIC	LOGIC	TERMS	TERM WEIGHT
	4 5	6 7	8	9	10	DO NOT KP	11 DO NOT KEYPUNCH ASTERISK 77	78 80

FIGURE 12

CA CONDENSATES CITATION

CA CONDENSATES
DIGEST NO 106543B VOL. 73 NUMBER 21 REF WT. 000000
(1) JABAA4 0033 0002 70 0380 0009
(2) J. APPL. BACTERIOL.
(3) DEGRADATION OF NUCLEIC ACIDS BY CYTOPHAGA JOHNSONII.==
(4) GREAVES MP, VAUGHAN D, WEBLEY DM.
(5) (MACAULAY INST. SOIL RES., ABERDEEN, SCOT.).

FIGURE 13

CA CONDENSATES CITATION ILLUSTRATING APPLIED TEXT WORDS

CA CONDENSATES
DIGEST NO 106976V VOL. 73 NUMBER 21 REF WT. 000000
(1) HICHAU 0022 0004 70 0337 0046
(2) HISTOCHEMIE
(3) ZYMOGRAPHIC DEMONSTRATION OF LACTATE AND MALATE DEHYDROGE
NASES ISOENZYMES IN THE RODENT SALIVARY GLANDS.==
(4) YOSHIMURA Y, KAWANO T, KUROI M, MORISHITA M, MORI M, KAWA
KATSU K.
(5) (DENT. SCH., OSAKA UNIV., OSAKA, JAPAN).
LACTATE DEHYDROGENASES SALIVARY GLANDS
MALATE DEHYDROGENASES SALIVARY GLANDS
ELECTROPHORESIS DEHYDROGENASES SALIVARY GLANDS
SALIVARY GLANDS DEHYDROGENASES RODENTS
DEHYDROGENASES RODENTS SALIVARY GLANDS

Authors

Authors are coded in Column 9 as A, and the author's name is retrieved regardless of the number of individual authors listed for a single citation. An author search is highly successful in a Chemical Titles search and only partially successful in Chemical Abstracts. For either service it is best to list the author's last name followed by a space and two initials. However, in a CT search, it must read SINGH CM, while in CA it should be written SINGH C M (punctuation is never used).

Coden

Coden is a relatively new standard designed by the American Society for Testing and Materials (3) which describes accurately a specific journal title by use of a unique alpha/numeric code. Figure 14 gives a few specific examples of typical journal coden and is a copy of the transparency used to explain the coden field. As indicated in Figure 14, any alpha or numeric character separated by a hyphen should be searched as a solid string of characters. In our seminars we make a point of citing the exact campus location of this important Coden reference set.

CAS has prepared its own patent coden (Figure 15) which is based on the first two letters of the country of origin followed by the three alpha symbols XXA. Therefore, a French patent would be listed as FRXXA while the United States is listed as USXXA. This search has been useful as an identification tool as well as a search method.

CAS has also provided a Coden type symbol "BOOKA" which is used in CA Condensates on the Coden line to indicate a full book. Individual chapters of that book when on a specific monographic subject and by specific authors are again cited as the more familiar journal article type reference. Coden for journals and patents work equally well in a CT search. BOOKA is not included on the CT tapes but more importantly neither is the short title line following the Coden citation.

The symbol C representing the Coden search field should be entered in the Search Profile Form (Figure 11) in column 9 under "type". The actual alpha/numeric specific journal designation should be entered on the lines designated as columns 11-77.

---

[3] Coden for Periodical Titles. Volume I. Periodical Titles by Coden, Nonperiodical Titles Deleted Coden; Volume II. Periodical Titles by Title. American Society for Testing and Materials, Philadelphia, 1966. Supplements, 1968 and 1969.

FIGURE 14

CODEN

PERIODICAL-TITLES ALPHABETICAL BY TITLE

American Journal of Anatomy	AJAN-A
American Journal of Botany	AJBO-A
American Journal of Cancer	AJCA-A
American Journal of Cardiology	AJCD-A
American Journal of Veterinary Research	AJVR-A

Coden may appear in the Coden for Periodical Titles or on the journal using both alpha and numeric characters. These are usually separated by a hyphen but they should be searched as a solid string of characters.

Journal of Organic Chemistry

JOCE-A=JOCEA

Coden can be searched in any truncation mode but we suggest the use of mode 2.

FIGURE 15

PATENT CODEN AND ABBREVIATIONS OF COUNTRIES

COUNTRY OF ORIGIN	ABBREVIATION AS USED IN ABSTRACT	SPECIAL CODEN FOR PATENTS
Argentina	Argentine	ARXXA
Australia	Australian	ALXXA
Austria	Austrian	AUXXA
Belgium	Belg.	BEXXA
Britain	Brit.	BRXXA
Britain	Brit. (amended)	BSXXA
Canada	Can.	CAXXA
China	Chinese (Taiwan)	CTXXA
Czechoslovakia	Czech.	CZXXA
Denmark	Dan.	DAXXA
Finland	Finn.	FIXXA
France	Fr.	FRXXA
Germany	Ger.	DEXXA
Germany (East)	Ger. (East)	GEXXA
Hungary	Hung.	HUXXA
India	Indian	INXXA
Israel	Israeli	ISXXA
Italy	Ital.	ITXXA
Japan	Japan	JAXXA
Mexico	Mex.	MEXXA
Netherlands	Neth. Appl.	NAXXA
Netherlands	Neth.	NEXXA
Norway	Norweg.	NOXXA
Poland	Pol.	POXXA
South Africa	S. African	SAXXA
Spain	Span.	SPXXA
Sweden	Swedish	SSXXA
Switzerland	Swiss	SWXXA
Russia	U.S.S.R.	URXXA
United States	U.S.	USXXA
United States	U.S. Reissue	UUXXA

From: Preparation of Search Profiles, A Guide to the Effective Use of CAS Search Services on Magnetic Tape, Chemical Abstracting Service, Columbus, Ohio, 1968, p.41.

Title and Text Words

Title and Text Words are characterized in Column 9 as T since the same word might be found in either source. CT tapes only have title words while CA tapes contain keywords applied to designate primary abstract concepts although, of course, the complete abstract does not appear in the printout. Figure 13 illustrates how text words are applied to a CA Condensates tape citation at Chemical Abstracting Service. These added phrases appear following line (5) as previously illustrated in Figure 12.

Since a term-match system is used as the basic tape design, keywords must be very carefully and critically selected. With this fact constantly in mind, one must proceed to teach the fundamental philosophy and methods for analyzing problems when they arise. The text-word concept is by its very nature the most difficult to comprehend and most confusing to the user.

Certain principles must be carefully explained. In both CA and CT all words which might be used synonymously by colleagues throughout the world should be listed as title or text words. Spelling of the same word which varies in English usage must be listed in both forms, i.e., SULFUR and SULPHUR.

There are some words in the broad concept of chemistry which carry variant multi-disciplinary meanings. This latter problem can be described by defining the word CEMENT in the context of a structural engineer and a dentist. This problem can be somewhat eliminated in CA Condensates by encouraging the use of a profile search on either the "odd" or "even" numbered tapes. Figures 16 and 17 list selected subject headings and are used as graphic remainders to the user that the odd issues represent sections 1-34 which contain the biological and organic chemistry, while the even issues index sections 35-80 representing the areas of applied and physical chemistry.

Abbreviations should also be carefully explained using appropriate examples. The CAS Preparation of Search Profiles contains two pages of standard abbreviations, but is by no means complete. This manual stipulates that "plurals of noun abbreviations are formed by adding 's' to the singular abbreviations except when a single abbreviation is designated to show both singular and plural forms", and states "Words formed by adding prefixes to words normally abbreviated are also abbreviated, as microchem . . .". In addition, English units of weights and measures are abbreviated. Therefore, not being absolutely positive of the use of all abbreviations as either title or text words, one is sometimes far ahead by listing both, particularly in CT where only exact title listing is available. For example, in CT a search based strictly on "RNA" would not produce the same hits as "ribonucleic acid" nor would a search for "nitrogen" produce hits where only the

FIGURE 16
CHEMICAL ABSTRACTS (ODD)
Sections 1-34

BIOCHEMISTRY SECTIONS
- 3. Enzymes
- 5. Radiation Biochemistry
- 7. Plant Biochemistry
- 8. Microbial Biochemistry
- 9. Nonmammalian Biochemistry
- 11. Mammalian Biochemistry
- 12. Mammalian Pathological Biochemistry
- 15. Pharmacodynamics
- 20. Fertilizers, Soils, and Plant Nutrition

ORGANIC CHEMISTRY SECTIONS
- 22. Physical Organic Chemistry
- 27. Heterocyclic Compounds
- 29. Organometallic and Organometalloidal Compounds
- 32. Steroids
- 34. Synthesis of Amino Acids, Peptides and Proteins

FIGURE 17
CHEMICAL ABSTRACTS (EVEN)
Sections 35-80

MACROMOLECULAR CHEMISTRY SECTIONS
- 35. Synthetic High Polymers
- 38. Elastomers, Including Natural Rubber

APPLIED CHEMISTRY AND CHEMICAL ENGINEERING SECTIONS
- 53. Mineralogical and Geological Chemistry
- 59. Air Pollution and Industrial Hygiene
- 61. Water
- 64. Pharmaceutical Analysis

PHYSICAL AND ANALYTICAL CHEMISTRY SECTIONS
- 68. Phase Equilibriums, Chemical Equilibriums and Solutions
- 74. Radiation Chemistry, Photochemistry, and Photographic Processes
- 76. Nuclear Technology
- 79. Inorganic Analytical Chemistry
- 80. Organic Analytical Chemistry

symbol "N" appeared in the title. All of these facts were carefully explained in seminars; however those users preparing profiles with only the CAS provided pamphlets were at great disadvantage.

Chemical Abstracting Services recognized the problem and announced a keywording policy change effective with July 1970 CA Condensates and its corresponding issue (however, the announcement letter was dated October 20, 1970 and received two days later). The new policy eliminates the atomic and radical symbols in the Keyword Subject Index and instead spells out the names. This solves part of the problem in that we know that "Zn" will never be used as a tape keyword but will always be listed as "Zinc", and that "NaCl" will no longer be used as the simple formula for salt in the Keyword Subject List. There letter states:

"This change in Chemical Abstracts Service policy was made so that CA Condensates subscribers may use the tapes more effectively and economically by achieving consistency in element and radical name handling in the Keyword Subject Index. Studies show that while this practice will add some additional [sic] verbiage to the printed Keyword Subject Index, it will not do so appreciably."

It still appears, however, that both symbols and full spelling would be necessary in profiling according to normal terminology used in document titles since CAS does not normally appear to duplicate title words in the keyword list.

Even though these and other problems have resulted in considerable study, the search by title and keyword works extremely well unless complicated chemical systems are involved.

DESCRIPTION OF SEARCH STRATEGY

When a potential user has determined that a customized literature search will provide the help he needs, he begins the preparation of his profile. There are four primary areas of consideration:

I. The Statement of the Question

II. Identification of the Concepts

III. Expansion of the Concepts

IV. Refinements

We will follow this procedure through by illustrating an actual working profile. The question as stated is Combustion of Metals. The major concepts are identified and then further expanded as shown below:

COMBUSTION	METALS
IGNIT FLAME BURN	MAGNESIUM ZINC COPPER IRON MG ZN CU FE

The expansion entries shown above are examples selected from the total range of expansion concepts listed on the actual profile form (Figure 18). It is at the concepts expansion level that the user must decide the degree of precision he requires in retrieving relevant information. In our example, it was necessary to decide exactly which metals were pertinent to the researcher's particular problem. It is also at this point that the questioner must consider carefully the choice of terminology of his field and even more importantly the terminology of his foreign colleagues.

Before continuing with a discussion of the final area of Refinements, the reader should be aware of the usefulness of the Search Guide provided by Chemical Abstracting Service. This publication was prepared to "aid the questioner in accomplishing this expansion in the formulation of his search strategy". We have indeed found it useful as a guide to technical search strategy in addition to the concept expansion technique. A sample page (Figure 19) from the Search Guide is included to support the technique discussion. The publication is arranged alphabetically by terms selected by CAS scientists based on usage in the literature. The codes remind researchers of additional terms to be considered in expanding the question concepts. The Search Guide format is as follows:

MAIN TERM ——→ LIPEMIA

BT DISEASES ——→ BT - Broad Term
NT HYPERLIPEMIA ——→ NT - Narrow Term
RT ANTILEPEMIC AGENTS
RT BLOOD
RT FAT
RT FATTY EMULSION ——→ RT - Cross Reference Term
SY LIPAEMIA
SY LIPIDEMIA
SY LIPOIDEMIA ——→ SY or S* - Synonyms

"The symbol S* is used to cross refer all other synonyms to the one under which the set of associated terms is listed."

FIGURE 18. PREPARED PROFILE

SEARCH PROFILE FORM

Return to:
DIANE BENNETT
GENERAL LIBRARY, ROOM 204
UNIVERSITY OF CALIFORNIA
RIVERSIDE, CALIFORNIA

REQUESTOR: J. Byrne/J. Verity
AFFILIATION: Lawrence Radiation Laboratory
ADDRESS: P. O. Box 808
Livermore, Calif. 94550

HEADER CARD

QUESTION NUMBER (1–3)	OUTPUT INDICATORS (8–11)	QUESTION WEIGHT (12–17)
022	1000	

QUESTION: COMBUSTION OF METALS

CA CONDENSATES (SDI)

DETAIL CARD

DUPLICATE	PARAM NO. (4–5)	NO. OF PARAM (6–7)	MODE (8)	TYPE (9)	LOGIC (10)	LOGIC (DO NOT KP)	TERMS (11 – DO NOT KEYPUNCH ASTERISK – 77)	TERM WEIGHT (78–80)
	01	03	4	T	Ø		*CØMBUSTIØN*	
		DUPLICATE					*IGNIT*	
							FLAME	
							BURN	
	02	2					METAL*	
							MAGNESIUM*	
							ZINC*	
							CADMIUM*	
							CØPPER*	
							URANIUM*	
							BERYLLIUM*	
							TITANIUM*	
							BØRØN*	
							ZIRCØNIUM*	
							CERIUM*	
							MISCHMETAL*	
							ALUMINUM*	
							SØDIUM*	
							ALKALI*	
							PHØSPHØRUS*	
							BARIUM*	

FIGURE 18 (Continued)

DUPLICATE	PARAM NO. (4 5)	NO. OF PARAM (6 7)	MODE (8)	TYPE (9)	LOGIC (10)	LOGIC (DO NOT KP)	TERMS (11 DO NOT KEYPUNCH ASTERISK 77)	TERM WEIGHT (78 80)
		DUPLICATE					CALCIUM*	
							IRØN*	
			3				MG	
							ZN	
							CD	
							CU	
							U	
							BE	
							TI	
							B	
	02	03	3	T	Ø		ZR	
							CE	
							AL	
							NA	
							P	
							BA	
							CA	
							FE	
	03		2	T	N		PRØØF*	
							RESISTAN*	
							RETARDAN*	
							FLAMEPRØØF*	
							PHØTØMETR*	
							SPECTRØ*	
							ENGINE*	
							CELLULØS*	

FIGURE 19

SEARCH GUIDE

LININGS
RT EPITHELIUM
RT PLATING
RT PROTECTION
RT REFRACTORIES

LINKAGES
RT FITTINGS
RT JOINTS (MECHANICAL)

LINKAGES (CHEMICAL)
S* BONDS (CHEMICAL)
SY BONDS (CHEMICAL)

LINOLEUM
RT LINSEED OIL

LINSEED
SY FLAXSEED

LINSEED OIL
BT OILS
BT VEGETABLE OILS
RT LINOLEUM
RT PAINT
RT PRINTING INK
RT STAND OIL
RT TUNG OIL
RT VARNISH

LINTER
S* COTTON LINTER

LIPAEMIA
S* LIPEMIA

LIPASE
BT ENZYMES
NT PANCREATIC LIPASE
RT LIPOPROTEIN

LIPEMIA
BT DISEASES
NT HYPERLIPEMIA
RT ANTILIPEMIC AGENTS
RT BLOOD
RT FAT
RT FATTY EMULSION
SY LIPAEMIA
SY LIPIDEMIA
SY LIPOIDEMIA

LIPHOPHILIA
RT LIPOPHIL

LIPIDE
S* LIPIDS

LIPIDEMIA
S* LIPEMIA

LIPIDS
NT BODY FAT
NT CAROTENE
NT CEREBROSIDES
NT FATS
NT FATTY ACIDS
NT GLYCEROLIPID
NT GLYCO LIPID
NT PHOSOLIPIDS
NT PHOSPHATIDES
NT STEROIDS
NT STEROLS
NT TERPENE

LIPIDS
RT ANTILIPEMIC AGENTS
RT CARBOHYDRATES
RT CELL (BIOLOGICAL)
RT FAT
RT FATTY ACIDS
RT HYPOLIPEMICS
RT LIPOGENESIS
RT LIPOPHYL
RT LIPOPROTEIN
RT PROTEINS
RT WAX
SY LIPIDE
SY LIPOID

LIPIN
RT MYELIN

LIPOGENESIS
BT BIOSYNTHESIS
BT REACTIONS (CHEMICAL)
RT BODY FAT
RT FAT
RT LIPIDS

LIPOID
S* LIPIDS

LIPOIDEMIA
S* LIPEMIA
SY LIPEMIA

LIPOLYSIS
BT DECOMPOSITION
BT DEGRADATION
BT REACTIONS (CHEMICAL)
RT HYDROLYSIS

LIPOPHIL
RT LIPHOPHILIA
RT LIPIDS
S* LIPOPHYL

LIPOPHILIA
RT LIPOPHYL

LIPOPHYL
SY LIPOPHIL

LIPOPOLYSACCHARIDES
BT POLYSACCHARIDES

LIPOPROTEIN
BT PROTEINS
RT LIPASE
RT LIPIDS
RT LIPOPROTEINASE
RT LIPOVITELLENIN
RT LIPOVITELLIN
RT PROTEOLIPIDE

LIPOPROTEINASE
BT ENZYMES
RT LIPOPROTEIN

LIPOVITELLENIN
RT LIPOPROTEIN

LIPOVITELLIN
RT LIPOPROTEIN

LIPOXIDASE
BT ENZYMES
RT DOUBLE BONDS

Either terms or phrases may be used as search terms and they need not be confined to terms found in the CAS Search Guide.

Returning to the idea of concept refinement, we find this last step accomplished by applying a certain technical strategy to the selected terms through the use of coding and logic. The refinement process normally continues for a few weeks. During this period the retrieved material is evaluated and the search profile is often revised.

Parameters

Parameters as a device for coding search strategy, are discussed first. A parameter is simply a subset of expanded terms within a single profile. Therefore, in our example of a prepared profile (Figure 18), columns 4 and 5 are reserved for the parameter number. The first parameter number is indicated by 01 and all terms selected as of equal interest within one concept are added to that parameter number. The second concept is identified as the second parameter, 02, etc. One can also think of the first parameter as identifying the primary concept of the profile question and each term, abbreviation, synonym, etc., so considered should be coded as parameter 01 in column 4 and 5. Parameter 02 would then contain modifying terms. Columns 6 and 7 should give the total number of parameters in the search question (03 being used for Figure 18). A search may be done in one of more parameters, but we suggest that a single question be limited to three parameters or less for the best results.

Logic

Logic relationships among the terms in a search profile are used to inform the computer of the user's intent. Parameters when properly used express the required relationships. Three types of logic (Or, And, Not) are applicable to the Chemical Abstracting Service Programs. "Or" logic is the most frequently used and must be coded as Ø in Column 10 of the search profile form (Figure 18).

All terms within a single parameter are interrelated by "Or" logic. The presence of any one of these terms will then satisfy that particular parameter of the search profile. For example:

COMBUSTION

[or]

IGNIT

[or]

FLAME

[or]

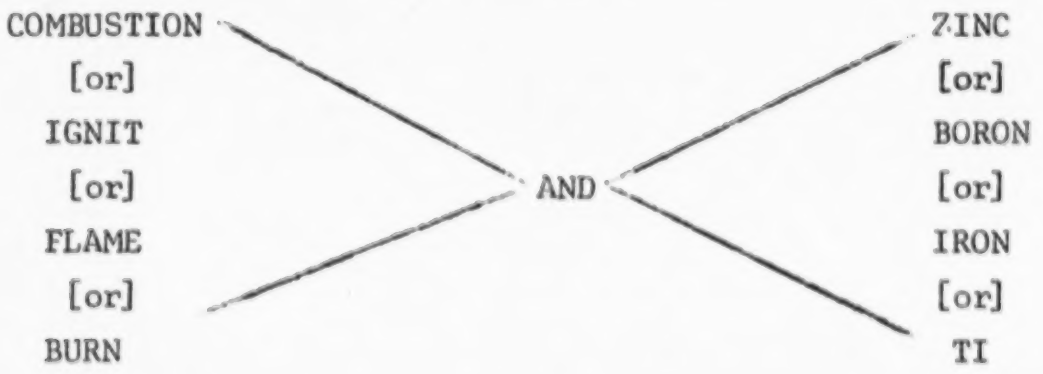
BURN

Therefore, if a one-parameter search is undertaken, the system will accept any citation in which one of these terms appears regardless of its application.

The second type of logic is "And" logic. "And" logic is always implied between the parameters of a single search profile. Being an automatically implied logic, it is not coded on the profile form. This logic form requires that a reference must satisfy all of the parameters before that reference will be retrieved. In our example, the second parameter contains in part:

ZINC
[or]
BORON
[or]
IRON
[or]
TI

With the "Or" logic existing between all terms of a single parameter, we have now established a second logical requirement that must be satisfied between parameters before a reference will be retrieved. For example:



A combination of any term from the first parameter AND any term from the second parameter will retrieve a citation. It should be quite clear that the greater the number of parameters the greater the amount of restriction and hence the fewer number of citations capable of satisfying the logic.

"Not" logic is the third and final type present in the CAS programs. This logic provides the capability to eliminate unwanted terms regardless of how the reference may otherwise satisfy the profile requirements. For example, "Not" logic was applied to the above example by including

NOT:

PROOF

[or]

FLAMEPROOF

[or]

ENGINE

We have hereby created a three parameter search primarily through the use of logic in the evaluation and expansion of our stated search question. All "Not" logic terms must appear in a single parameter and we prefer this is to be the final parameter of the profile. "Not" logic must be coded in Column 10 as N.

To complete the description of parameters and logic, the user should be aware that an additional way of selecting terminology and expanding a concept is by recognizing low frequency and high frequency terms. For this purpose CAS prepared the manual Chemical Titles Word Frequency List. One must remember that this is merely a guide representing the frequency of terms used in approximately twenty issues of Chemical Titles. The frequency is indicated by a different alpha code for stated number of occurrences. For example, A represents 1 occurrence in 20 issues while C indicates 11 to 100 and F equals over 1,000. It follows then that "Terms having high frequency of occurrence have little discriminating power and are likely to retrieve large numbers of references. Conversely, low frequency terms have much greater discriminating power, but tend to produce low yields of answers. Low frequency terms can be placed in a single parameter with the use of "Or" logic in order to increase the number of references." Low frequency terms can also be placed in a second parameter to modify high frequency words used in the first parameter. We have found that a profile containing four or less parameters provides greater user satisfaction. Relatively short precise profiles give a far greater percentage of relevant citations. The more parameters present, the greater the reduction in possible "hits" (citations). It is the user's responsibility to decide the number of relevant items he is willing to miss and/or the extraneous items he is willing to pay for.

Term Truncation

Term truncation is another technical device available in search strategy. It is designed to enable expansion of word fragments to include other combinations of characters preceding or following the

shortened term or fragmentation. Figure 20, drawn from a CAS manual, describes how this system works. According to CAS, "The asterisk denotes that a blank space or any character may occupy that position." Individuals preparing profile forms are requested to include the asterisk with the term or term fragment even though the mode column number supplies the information to the program. This is an important double check of a very versatile system which can also be dangerous. Mode 1 accepts any prefix or characters preceding the term; mode 2 accepts any suffix or characters following the term; mode 3 accepts only the term itself when bounded on both sides by a blank; mode 4 accepts the term and any or all additional characters. The search mode numbers must be correctly coded in column 8 of the profile forms.

Figure 21 illustrates how the term truncation mode works for the word "polymer". Term truncation is a very powerful tool which allows you to easily expand words, eliminating the need to list each acceptable form. There are many dangers involved in this system and it must be carefully used. Sometimes short terms can also be a fragment of a totally different high frequency word and as such are retrieved. For example, *MICA* in mode 4 is a perfectly acceptable search concept. However, it will retrieve all references to cheMICAL. A *HEME* search in mode 4 will retrieve all references where scHEMEs are mentioned. In addition to these errors, false hits can be obtained in even more obvious ways. GAS* searched in mode 2 to include GASES and GASEOUS would also extract references for GASkets, GASoline, GAStropod, GAStric and other similar medical terms. We suggest the use of the CAS Search Guide and very careful analysis of fragmented search terms.

Before leaving word fragmentation completely, we must point out that in CT word fragments are used in a different concept having nothing to do with truncation mode.

The hard copy publication of Chemical Titles is produced by the KWIC (Keyword-in-Context) indexing method using only titles as they appear in the primary journals. To properly index these titles by subject, some common scientific words have been fragmented so as to be machine indexed in more than one place. For example:

Carbon Di Oxide

Carbon Tetra Chloride

Chole Sterol

Thi Amine

Words are identically fragmented on the CT magnetic tapes; therefore, they must be properly fragmented on the Search Profile Form and properly fragmentated in keypunching if the search is conducted on CT. It is not necessary to do fragment terms when using CA Condensates unless you can use a truncation mode and are interested in, for example, all

FIGURE 20
TERM TRUNCATION MODE - EXPLANATION

SEARCH MODE	SYMBOLIC REPRESENTATION	EXPLANATION
1	*term	Accepts any combination of characters preceding the term.
2	term*	Accepts any combination of characters following the term.
3	term	Accepts term only if bounded by blank spaces or non-alphabetic characters.
4	*term*	Accepts the term with any string of characters either preceding or succeeding the term.

FIGURE 21
TERM TRUNCATION MODE - EXAMPLE

Mode 1

*Polymer = Copolymer
Biopolymer

Mode 2

Polymer* = Polymerase
Polymers
Polymerization

Mode 3

Polymer = Polymer

Mode 4

Polymer = All combinations above, plus such terms as:
Photopolymers
Copolymerization
Depolymerization
Pre-polymerization

of the amine compounds. If in doubt as to the possible fragmentation of the term you wish to use, we suggest that you check the word in any issue of Chemical Titles and follow that form.

Term Weight

Term weight is a numerical value which may be assigned to each search term related to the value of that term to the question. These weights may be positive or negative and applied to all parameters except those containing "Not" logic. Ideally the total cumulation of term weights when added or subtracted must equal or be higher than the threshold weight applied to the total search profile question before a citation would be received. The weighting factor, however, is considered only after all parameters have been satisfied. Term weighting factors are coded in columns 78 to 80 and high-order zeros should be coded. If a positive weighting factor of 25 were being applied it would be coded as 025. If a negative weighting factor of -10 were applied, the minus sign would appear in column 78 and the ten in columns 79 and 80. The total cumulated search profile weight is coded in the header card, columns 12 to 17, where high-order zeros are also coded. If the total weight assigned is 85, this is written 000085.

Negative weighting can be highly useful and should often be applied to undesirable concepts in place of "Not" logic. As an example, consider a search involving the effects of certain insecticide residues on soil or plant types. One could easily eliminate certain soil or plant types or even certain chemical compounds by name if they were wholly undesirable through the use of "Not" logic. However, many journal titles include "fungicides" and "herbicides" with insecticides, and some might include both desired and undesired plant family or soil types. If "Not" logic is used, it will override all other logic and weighting factors and the article will not be retrieved. On the other hand, the use of negative weighting permits the article to be retrieved if the cumulated positive term weights minus the negative term weights still equals or exceeds the total question weight.

Citations are printed on the output forms in decreasing order of total reference weight. This order of arrangement is automatically established by the CA system which matches each term in the search profile and cumulates them into a completed reference weight which is then compared to the preselected question weight. The retrieved file is thus produced on the printout in decreasing order, which theorectically places the most relevant items near the beginning of the output.

Columns 78 to 80 are used to indicate the individual weighting factor applied to each specific term. The sum of the terms matched (through which a citation is retrieved) becomes the reference weight. If the total reference weight matches or exceeds the question weight precoded in the header card (columns 12 to 17), the reference will be

retrieved and printed in decreasing order. For the search profile used in Figure 22, a sample page of output (Figure 23) illustrates the placement of the various weighting factors and the decreasing order of retrieval.

In teaching search procedures, it is also very important to emphasize for some patrons the kinds of borderline, life science information which Chemical Abstracting Service has intentionally not included. They do not abstract material on morphology studies, taxonomy reports, or anatomy studies. They abstract these only when the document refers to the chemical action of a specific compound as it relates to the reported diagnostic studies, clinical studies, immunological studies, and pathological studies. Constraints such as these are discussed during a seminar and visual examples are presented.

FIGURE 22

WEIGHTING FACTOR CODING ON SEARCH PROFILE FORM

HEADER CARD			QUESTION: Chemical Analysis of Sea Water
QUESTION NUMBER	OUTPUT INDICATORS	QUESTION WEIGHT	
1 3	8 11	12 17	
0,6,2	1,0,0,0	0,0,0,1,4,0	

DETAIL CARD

DUPLICATE	PARAM NO.	NO. OF PARAM	MODE	TYPE	LOGIC	LOGIC	TERMS	TERM WEIGHT
	4 5	6 7	8	9	10	DO NOT KP	11 DO NOT KEYPUNCH ASTERISK 77	78 80
	0,1	0,2	3	T	Ø	ØR	SEA	0,5,5
	0,1		3	T	Ø	ØR	WATER	0,5,0
	0,1		3	T	Ø	ØR	MARINE	0,5,5
	0,2		3	T	Ø	ØR	ANALYSIS	0,4,0
	0,2		2	T	Ø	ØR	TECHNIQUE*	0,1,0
	0,2		3	T	Ø	ØR	COMPOSITION	0,4,0
	0,2		2	T	Ø	ØR	CHEMI*	0,4,0

FIGURE 23

WEIGHTING FACTOR CODING ON
PRINT-OUT FORM

NAME QUESTION NO 062 DATE 07/05/70
THRESHOLD WEIGHT 000140

CHEMICAL ANALYSIS OF SEA WATER

CA CONDENSATES
DIGEST NO 002072F VOL. 73 NUMBER 01 REF WT. 000260
FCONAA 27,, 0000 69 0322 0039
FORTSCHR. CHEM. ORG. NATURST.
CHEMISTRY OF SOME TOXINS ISOLATED FROM MARINE ORGANISMS.==
SCHEUER PJ.,
(DEP. OF CHEM., UNIV. OF HAWAII, HONOLULU, HAWAII).

CA CONDENSATES
DIGEST NO 001401A VOL. 73 NUMBER 01 REF WT. 000190
NUDIAI 0011 0004 69 0280 0092
NUTR. DIETA
EFFECT OF WATER RESTRICTION ON THE STORAGE OF FAT AND ITS FATTY ACID COMPOSITION.==
DU PAN R M, ROSSI J, MOIOLA A, KOECHLI B.,
(SWITZ.).

CA CONDENSATES
DIGEST NO 001271H VOL. 73 NUMBER 01 REF WT. 000150
CHDDAT 0270 0006 70 0828 0030
C. R. ACAD. SCI., SER. D
MODE OF FORMATION, ULTRASTRUCTURE, AND CHEMICAL NATURE OF MICRONUCLEOLI OF A PARTICULAR TYPE DURING OVOGENESIS IN PARACENTROTUS LIVIDUS.==
SANCHEZ S.
(LAB. HISTO. EMBRYOL., FAC. MED., MONTPELLIER, FR.).

CA CONDENSATES
DIGEST NO 001486G VOL. 73 NUMBER 01 REF WT. 000140
PDANAN 0007 0004 68 0291 0006
PROD. ANIM.
RESEARCHES ON WATER BUFFALO MILK. IV. CARBOHYDRATE COMPOSITION OF CASEIN.==
ALBONICO F, MINCIONE B, ESPOSITO AM.
(FAC. AGRAR., UNIV. NAPOLI, PORTICI, ITALY).

VI. EVALUATION WORKSHOPS AND INTERVIEWS

Follow-up evaluation workshops were originally held primarily with small groups of reference librarians whom we anticipated would have the responsibility for all additional work with the users. The level of help varied from campus to campus depending upon the interest, subject and technological aptitude of the reference staff. This was by no means the major problem; we found that for the most part the user would prefer to analyze his own output which was indeed much more satisfactory and an exercise by which he gained invaluable systems knowledge to be used with additional profiles or new data bases. Since our staff organization had not increased, individual attention, generally following the course outline of Figure 24, was provided only upon individual user request. This outline was prepared for the evaluation sessions of the Center for Information Services seminars given at UCLA. Through workshop methods each student evaluated his own printout following a brief lecture in conjunction with the CAS manuals and other hard copy publications, visual aids and individual problem-solving techniques. Figure 24 is quite self-explanatory in that normal types of errors are studied, and a straightforward approach is used in the initial evaluation.

We have tried and analyzed to some extent the validity of the personal interview and indirect studies approach to the secondary evaluation methods. We have not used the questionnaire although we do use a User Feedback Card (Figure 25) which is mailed with each printout until the user is fully satisfied with his results. Each quarter an SDI Update Card (Figure 26) and a search profile printout (see following example) are sent to the user for profile updating and additional feedback.

FIGURE 24

INSTITUTE OF LIBRARY RESEARCH
CENTER FOR INFORMATION SERVICES

Seminar on Mechanized Information Service
in the University Library

Evaluation Workshop

Input and output must be completely anlyzed repeatedly until the maximum success is achieved. The Librarian or Analyst should first check obvious error sources.

I. Initial Evaluation

A. Possible types of errors
   1. Relevancy-keyword error
   2. Coding or keypunch error
   3. Fragmentation or truncation error
   4. Section or issue error

B. Analysis of output
   1. Type of error
   2. Parameter problems
   3. Weighting problems
   4. What is retrieved versus what is retrievable?

C. Analysis of input
   1. What happened?
   2. Why did it happen?

D. A guide for the initial evaluation
   1. Check all hits quickly and indicate only those whose relevance you question.
   2. Check printout of profile for keypunch errors.
   3. Underline words in each parameter which pulled your hit. Use different color, one, two, and/or three straight lines, or one wavy and one straight.
   4. Look for words pulled by truncation mode. Use Search Guide.

FIGURE 24 (Continued)

5. Go to hard copy.
   Check forms of words used:
   fragmentation or abbreviations
   hyphen and/or other punctuation
   spelling
   Check index and/or section of interest to see if items were missed and recheck profile to see why.
6. Check Search Guide again for other terms to be included.
7. Use Word Frequency List to help broaden or narrow the terms.

II. Secondary Evaluation

A. Questions, problems and discussions to retrieve user feedback.
   1. Personal interview
      a. May be structured or unstructured. Early interviews are generally unstructured while the interviewer feels out what questions need to be asked.
      b. Highly individual. Perhaps one should attempt to become more structured as you both become more aware of the problem.
      c. Most time consuming
      d. Most productive
      e. Best for public relations

   2. Indirect Studies
      a. Examination of the data itself
         1. Number of hits received
         2. Percentage of relevancy
         3. Percentage of retrieval vs. amount to be retrieved.
         4. Restudy of profile itself with structured questioning raised during initial evaluation.
      b. Sometimes such an analysis will lead to clear cut conclusions; in other cases it may be ambiguous and require further analysis or a combination of personal interview with already obtained analytic results.

   3. Questionnaire
      a. Rather shallow form of feedback and least effective method.
      b. No way of measuring the:
         1. Clarity of questions
         2. Receptivity of the questioner.
         3. Lower response through returns
         4. Considerable variation in time and effort spent in reply.

FIGURE 24 (Continued)

c. If this method is used, keep several important things in mind:
   1. What information are you seeking?
   2. What is the logical sequence of questions?
   3. Does the respondee have the necessary information to answer the questions?
   4. Is the questionnaire form easy to respond to?
   5. Recurring questionnaires designed to detect changes in attitude generally are not effective because most people will not report on typical behavior or will base one reply on the last.

FIGURE 25

USER FEEDBACK RECORD

NAME:

QUESTION NO.:

VOL. & NO.:

DATE RECEIVED:

TOTAL NO. OF HITS:
Relevant:
Irrelevant:

Signed _____

FIGURE 26

SDI UPDATE

_____ Quarter

NAME: _____	CAMPUS EXTENSION: _____
CAMPUS: _____	SERVICE: _____
FUND: _____	QUESTION NO.: _____
ACTION: Revision Attached _____ Drop Effective _____ Reason _____ _____	Send Printout By: Campus Mail ____ Airmail ____ Intercampus Bus _____ First Class Mail _____
Correct the Following: From: _____ To: _____ Run only on ODD / EVEN / ALL Numbers of CA	☐ I request consultation before writing/revising profile. ☐ I am delighted with results. ☐ I am satisfied with results.
Please Use Verso for Remarks or Suggestions	

The search profile printout (question 62) shown below was generated for the profile used in Figure 22. The parallel lines are drawn to represent the coding columns on the Search Profile Form.

SEARCH PROFILE PRINTOUT

```
062|  |  |10000001 40
062|01|02|3|T|0|SEA                    55
062|01|02|3|T|0|WATER                  50
062|01|02|3|T|0|MARINE                 55
062|02|02|3|T|0|ANALYSIS               40
062|02|02|2|T|0|TECHNIQUE              10
062|02|02|3|T|0|COMPOSITION            40
062|02|02|2|T|0|CHEMI                  40
```

We found it necessary to give most users additional help with the term weighting process. They seemed aware of terms which they used in a hard copy search, but were usually not familiar with synonyms or variant spelling. We tried to bend the user's thinking to relative values attributed to specific parts of his question and then list those in order of decreasing value.

We refer to this method as the calculation of weights by patron preferences, but other obvious benefits resulted. Help with the analysis of his early runs was sometimes given by preparing five or more column headings which were written down and each citation categorized under a heading column. For example, the analysis for a biochemically oriented physiologist wanting to recover information on several enzymes of glycolysis appeared as follows:

1	2	3
regulation		
	activity	
		structure.

The keywords from each citation were underlined and the word written in the corresponding column. It not only became quickly obvious to the patron which areas were best covered by his profile but gave him a guide for the application of term weights to that profile. This often led to greater understanding of the weaknesses and strengths of his profile and greatly increased revision capability.

A trial input/output evaluation by a biochemically oriented Ph.D. candidate was provided experimentally over a five week period as an alternative to the evaluation workshop. The function of this service

was to include revision of unsatisfactory profiles and provide help in the preparation of new profiles. The results of this test proved satisfactory and in fact helpful for the following reasons:

1. It permitted revision of profiles to user satisfaction with limited user involvement.
2. It aided in overcoming the initial fear of computer technology which some users were reluctant to attempt on their own.
3. It underlined the personal nature of an individual problem and the need for customized searching.
4. It permitted the development of a "case history" approach to satisfactory profile development.
5. It resulted in an accumulation of experience in preparing profiles and understanding tape configuration, and inconsistencies which resulted in a more objective and therefore more satisfactory interview.

VII. PRINTABLE DATA ELEMENTS AND FORMAT

As CT and CA Condensates vary somewhat in search capabilities and procedures, so also do they in print format and printable data elements.

CT provides only a reference to the primary journal citation and retrieves these references through KWIC indexing methods using only the author's title words as searchable subject terms. Therefore, the CT output data elements include only the:

Title of the Article

Author(s) Name

Primary Journal Reference

- Coden for Journal
- Volume Number
- Issue Number
- Year
- Pagination

CT Reference

The CT printout format as provided by the CAS program (Figure 27) presents all of the data print elements in readable form. Unfortunately, the CA Condensates printout format (Figure 28) did not prove to be so legible. However, all of the searchable data elements did appear even if somewhat more difficult to identify. The output data elements for CA Condensates include:

Chemical Abstract Reference
- Volume, Number, Abstract Number

Primary Journal Reference
- Coden, Year, Volume, Pages, Brief Title

Title of Article or Patent

Author(s) or Assignees

Location of Work
- Senior Author Address

FIGURE 27

CT PRINTOUT FORMAT (CAS)

ANSWERS TO CHEMICAL TITLES SEARCH INQUIRIES — PAGE 0001

(Tape Service Name)

QUESTION NUMBER 020 — QUESTION WEIGHT — (Reference Weight If Used)

(All Authors)

AJPHAP-0217-1708 CARRIERE S FRIBORG J 000000
INTRARENAL BLOOD FLOW AND P-AMINO HIPPURIC ACID
EXTRACTION DURING ANGIOTENSIN INFUSION.=
69 1708-15 CT6925

(Journal Coden / Volume No. / Beginning Page)

KLWOAZ-0047-1247 WERNING C SIEGENTHALER W 000000
RENIN - ANGIOTENSIN - ALDO STERONE SYSTEM VIEWED PATHO
PHYSIOLOGICALLY.=
69 1247-55 CT6925

(Print Title)

KLWOAZ-0047-1256 WERNING C BAUMANN K GYSLING E 000000
STIEL D VETTER W WEIGMANN P
SIEGENTHALER W
RENIN AND ALDO STERONE DURING IDIOPATHIC EDEMAS.=
69 1256-62 CT6925

KLWOAZ-0047-1263 KORZ R FISCHER F BEHN C 000000
RENIN-ANGIOTENSIN SYSTEM IN HYPERVOLMIA SIMULATED BY
IMMERSION.=
(Year) 69 1263-7 CT6925

(Beginning & Ending Pages)

(Chemical Titles Hard-Copy Ref. Volume & Issue)

0004 HITS FOR QUESTION NUMBER 020

FIGURE 28

CA CONDENSATES PRINTOUT FORMAT (CAS)

SEARCH RESULTS

Question Weight If Used

THRESHOLD WEIGHT--000020

QUESTION NUMBER 006 DIGEST NUMBER VOLUME + ISSUE REFERENCE WEIGHT

Journal Coden | Year | Abstract Number | Beginning & Ending Page Numbers | Of Hard-Copy Chemical Abstracts | If Used

065737D 70 15 000900

ABBIA4|0129|0002|69|0438|0046
ARCH. BIOCHEM. BIOPHYS.
INDUCTIVE EFFECT OF RICKETS ON GLYCOLYTIC ENZYMES OF RAT EPIPHYSEAL CARTILAGE AND ITS REVERSAL
MEYER WL, KUNIN AS,.
(COLL. OF MED., UNIV. OF VERMONT, BURLINGTON, VT.).

Volume Number | Issue Number

064997H 70 15 000600

JOCRAM|0039|0001|69|0075|0007
J. CHROMATOGR.
GAS-LIQUID CHROMATOGRAPHY OF ERGOCALCIFEROL AND CHOLECALCIFEROL IN NANOGRAM QUANTITIES.==
WILSON PW, LAWSON DE M, KODICEK E,.
(UNIV. CAMBRIDGE, CAMBRIDGE, ENGL.).

Senior Author Address | Title of Journal Article

065775M 70 15 000600

JCINAO|0048|0001|69|0084|0095
J. CLIN. INVEST.
EFFECT OF MIXED MICELLAR LIPID ON THE ABSORPTION OF CHOLESTEROL AND VITAMIN D3 INTO LYMPH.==
THOMPSON GR, OCKNER RK, ISSELBACHER KJ,.
(HARVARD MED. SCH., BOSTON, MASS.).

Brief Title

All Authors

071031S 70 16 000600

APPHAX|0025|0004|68|0435|0009
ACTA POL. PHARM.
STABILIZATION OF VITAMIN DRUGS. III. EFFECT OF STERILIZATION ON THE STABILITY OF INJECTION SO
NERLO H, PALAK W,.
(AKAD. MED., LUBLIN, POLAND).

VIII. COST ESTIMATES AND FUNDING

As previously mentioned, cost studies were not actively considered as a part of this project, which originated as a library service function of the Bio-Agricultural Library to accommodate the science clientele on the Riverside Campus.

During the first three months of running free experimental search profiles for the Ad Hoc Users' Committee, some cost analysis was performed based entirely upon the equipment available and computer rates at UCR in 1969. The cost per hit leveled off, at approximately 400 hits per run, at slightly less than $.20 per hit. Based on recorded statistics (Figure 29) and the anticipated number of user/profiles, we determined that our user charges would be based on a run/hit price even though all other existing services used some other factor. We are still using the run/hit pricing structure but have recently established an upper limit to number of hits charged.

The cost figures shown on Figure 30 cover the first two years of actual experimentation and operation at UCR. As, illustrated, the bulk of our income was taken from the UCR library budget and augmented during 1969/70 by funds from the Institute of Library Research NSF Grant (GN-827), the UCR Graduate Division Office, and Engineering Budget. The deficit, which dropped remarkably between the two years, was generously absorbed by the Computing Center. Costs rose considerably as profiles increased and as hourly salaries and computer fees also increased.

Since our project is entirely batch processing on overnight computer time, excessive computer speed is not of major importance to the turn-around time involved. However, CPU time and I/O counts are of major interest from a system viewpoint. To this end, the UCR programmer modifications to the CAS program has greatly speeded up the run, thereby reducing cost. As users and profiles increase, our information retrieval project should become self-supporting as an operational function. Numerous requests have been made to local campus administration for financial support, primarily for Research and Development, to no avail. We do have their strong moral support but with the excessive strains on local budgets, financial support has not been available.

FIGURE 29

SAMPLING OF COST ANALYSIS PER PROFILE

Type	Vol.	No.	No. of Hits	No. of Words	Cost of Entire Run	This Person's Share
CA	70	1	37	20/1	57.04	5.56
		2	5	20/1	58.72	3.00
		3	16	20/1	55.44	3.51
CT	69	1	15	20/1	65.44	
		2	18	20/1	50.32	
CA	70	1	155	17/1	57.04	24.58
		2	268	17/1	58.72	39.45
		3	89	17/1	55.44	19.06
CT	69	1	53	17/1	65.44	
		2	53	17/1	50.32	
CA	70	1	2	5/1	57.04	3.00
		2	0	5/1	58.72	3.00
		3	0	5/1	55.44	3.00
CT	69	1	2	5/1	65.44	
		2	0	5/1	50.32	
CA	70	1	2	18/3	57.04	3.00
		2	1	18/3	58.72	3.00
		3	2	18/3	55.44	3.00
CT	69	1	0	18/3	65.44	
		2	5	18/3	50.32	
CA	70	1	0	19/3	57.04	3.00
		2	1	19/3	58.72	3.00
		3	0	19/3	55.44	3.00
CT	69	1	0	19/3	65.44	
		2	4	19/3	50.32	
CA	70	1	48	23/1	57.04	7.70
		2	2	23/1	58.72	3.00
		3	42	23/1	55.44	9.05
CT	69	1	23		65.44	
		2	16		50.32	
CA	70	1	66	45/2	57.04	10.49
		2	20	45/2	58.72	3.00
		3	69	45/2	55.44	14.80
CT	69	1	40		65.44	
		2	35		50.32	

FIGURE 30
TWO YEAR COMPARATIVE INCOME AND EXPENDITURES

INFORMATION RETRIEVAL COMPARATIVE EXPENDITURES

	1968/69		1969/70	
SUBSCRIPTIONS				
CA Condensates	$4,000.00	1969 (annual)	$4,000.00	1970 (annual)
CT	1,700.00		1,700.00	
Total	$5,700.00		$5,700.00	
SUPPLIES				
Magnetic Tapes	411.60		1,774.00	
Airmail postage	62.66		434.28	
IBM Cards, etc.	2.20		2.20	
Travel	98.60		286.55	
Total	$575.06		$2,497.03	
COMPUTING CENTER	COST*	HOURS	COST*	HOURS
360/50 CPU Time	$2,711.16	13.29	$6,987.51	37.19
Programmer	1,676.91	257.00	5,736,86	782.00
Other	55.87	20.02	100.28	34.73
Total	$4,443.94	290.31	$11,924.65	853.92

* Expenditures represent primarily operational funds although a minimum of R&D is being accomplished.

COMPARATIVE INCOME

UCR Bio-Ag. 19900			
Books	$5,700.00	$9,500.00	
Staff	3,086.00	3,500.00	
S & E	707.80	618.23	
UCR Graduate Division		312.85	
Institute of Library Research (NSF)		2,000.00**	
Recharges			
UCR		1,844.00	
Other		1,188.05	
Other Outside Funds		1,000.00	Engineering
Total	$9,493.00	$19,963.13	
TOTAL EXPENDITURE	$10,719.00	$20,121.68	
TOTAL INCOME	9,493.80	19,963.13	
TOTAL DEFICIT	$ 1,225.20	$ 158.55	

** By December 1970 a total of $6000 was received from ILR.

IX. RESULTS AND CONCLUSIONS

"As the births of living creatures at first are ill-shapen, so are all innovations, which are the births of time."(4) Bacon could well have been writing about computerized innovations for his statement is equally true for Chemical Abstracting Service as an innovative tape service as it is for the University of California, Riverside, as an innovative supplier of that service. In time to come we could become known as ill-shapen idiots as the original innovations give way to far more sophisticated methods through experimentation and development. The ideas expressed here in no way underestimate the capabilities of CAS or downgrade results for UCR; instead, they concentrate on problems or concerns that may or amy not be of interest six months hence. Discussions of many problems relating to the configuration of the tapes themselves have taken place with CAS during the past three years. Communication has often been one-way and advice or explanation has not always been received. Certainly Chemical Abstracting Service is aware of the many recommendations we have made, and we hope most of them will be included in the new Standard Distribution Format.

CT, Chemical Titles tapes, being primarily for current awareness use, have only a short-term value and will probably not be retained indefinitely. The exception, of course, will be the set of CT tapes covering January 1962 through December 1968 (prior to the general publication of the CA Condensates base) since they can provide retrospective searching through almost ten years of chemical literature. Keyword-in-Context (KWIC) indexing does offer limitations in subject analysis, but fortunately all references covered by CT are repeated after a lengthy time-lag in the CA Condensates tapes with some additional subject analysis. With the CT tapes we have experienced very few major difficulties, none of which have been disastrous.

One problem has been in word fragmentation, which is not only somewhat difficult for a layman to understand, but has been on occasion inconsistent. A second problem pertains to both tape services, actually, and causes considerable consternation among clientele. CT, advertised to arrive approximately 35 days prior to the hard copy, has not proven entirely successful. CA monthly tapes also were expected to arrive 30 to 35 days prior to the hard copy. Records have shown that due to CAS

---

[4] Bacon, Francis. Of Innovations, an Essay.

equipment down-time, mail delivery discrepancies, etc., the average lead time for both services has been 14 days. Occasionally two tapes have been received together for the same service with one tape arriving almost in conjunction with its hard copy issue. This problem is more severe in CT because it is a current awareness media and some researchers are practically waiting at the door for their printout.

A third (and most unfortunate) problem with CT is the omission of certain print capabilities. It would have been exceedingly helpful to have a short title given in conjunction with the Coden as done for CA Condensates. This lack of prime information on the CT tape and subsequently on the user printout causes an unnecessary step in identifying the data to be performed either by the user or more usually the reference librarian.

Some incomplete tapes have been received where all references have not been included on the tape. We have suggested that a utility program to test the sequenc of digest numbers be implemented to detect such deficiencies.

AUTHORS

An author search works beautifully on CT but is only partially successful in CA Condensates. The CAS manual states in its preparation material that "Author terms must be searched in truncation mode 2 with a 1401 program, but can be searched in any mode with the 360 program". This is basically true for CT but we found that truncation mode 2 provided the best results on the 360 as well for CA since Chemical Abstracts is inconsistent in the indexing of author names.

In the CAS program, SINGH would produce SINGH C or SINGH C M, but SINGH C M as a profile term would not produce SINGH C. This is well and good as long as complete initials are always used in indexing, but the CAS manual does not explain that, even though the term SINGH would produce both SINGH C and SINGH C M, it would also retrieve:

SINGH B B
SINGH H
SINGH K N

all of which would be false hits. Inconsistencies also appear since CA sometimes indexes other irregularities such as:

Akhmedov Sh T

Bergel' son L D

Carta De-Angeli L

Chen S-M

De Robertis E M Jr.
Hsu H H-T
Huynh-Trong-Hieu
Marwell C E III
Sanchez Franco F
Sandoval Puerta J A

TITLE AND TEXT WORDS

Title and text words have been basically successful although in CA Condensates the selection of text keywords reproduced on the tape are meager and sometimes relationship with the abstract itself is not easily apparent. Chemical Abstracting Service does not state whether the terms used are selected by the author or the CAS abstractor. Other problems, previously touched upon, result in tape-produced discrepancies which provide false drops or nonretrieval. These tape-produced errors are primarily in misprints, misspelling or other keypunching errors. Some problems, however, are again caused by inconsistencies which are not adequately explained in the CAS manual, Preparation of Search Profiles. Terminology continues to be the prime pitfall of information retrieval and I hold little hope for effective change in the immediate future.

Keywords should always be systematically used in a fixed pattern approach and emphasized in the manual. Abbreviations are a problem when used as a keyword or as a title-term. The CAS manual does include a short list of CAS Standard Abbreviations as Appendix A on page 39 and 40, but it is by no means inclusive. Abbreviations in a truncated form are often used as text keywords applied at CAS. Example: CHEM* for chemistry and PHYS* for physics. Abbreviations have caused some other problems and standardization or a complete computer prepared list would be extremely useful.

WEIGHTING

Question weight, reference weight and term weight have worked extremely well with the CT tapes since no added terms are searched. It does not work equally effectively with CA Condensates. CAS admits that the results of term weighting can be inaccurate. The statement is made in the CAS manual that "The assignment of high weighting factors, relative to the question weight, to terms having a high frequency of occurrence in the information store being searched can result in the retrieval of file items with deceptively high weights. This results from programmed addition of the weighting factor for a term to the total weight of the item each time the term occurs in the item." And

therein lies the tale. The last sentence of that quotation is, in fact, extremely accurate while the remainder of the statement is in itself misleading. The high frequency of terms as analyzed in the CAS manual Chemical Titles Word Frequency List is not entirely responsible for the high frequency reference weight. As indicated in Figure 31, it is more often the result of added keywords from the abstract which appear to be supplied by the abstractors at Chemical Abstracting Service.

FIGURE 31

KEYWORDS APPLIED FROM ABSTRACT

CA CONDENSATES

DIGEST NO 106733P VOL. 73 NUMBER 21 REF WT. 000000
HAEMAX 0054 0007 69 0500 0010
HAEMATOLOGICA
ELECTROPHORETIC BEHAVIOR OF SOME ERYTHROCYTE PROTEINS (HEMOGLOBIN, LACTIC DEHYDROGENASE ISOENZYMES, MALIC DEHYDROGENASE, GLUTAMIC-OXALACETIC TRANSAMINASE, AND LEUCINE AMINOPEPTIDASE) IN CHICKEN EMBRYOS AND COCKERELS DURING DEVELOPMENT. = =
NOTARIO A, BOBBIO-PALLAVICINI E, MARCIANO S, DI MARCO M, BOBBIO-PALLAVICINI F.
(IST. CLIN. MED. GEN. TER. MED., UNIV. PAVIA, PAVIA, ITALY).

---

1. PROTEINS ERTHROCYTES EMBRYOS
2. ENZYMES EMBRYO MATURATION
3. EMBRYOS DEVELOPMENT
4. DEVELOPMENT EMBRYOS
5. HEMOGLOBIN EMBRYO MATURATION
6. ERYTHROCYTES PROTEINS EMBRYOS

CA CONDENSATES

DIGEST NO 106892Q VOL. 73 NUMBER 21 REF WT. 000000
JONUAI 0100 0008 70 0917 0024
J. NUTR.
INFLUENCE OF DIETARY LIPID ON LIPOGENESIS AND ON THE ACTIVITY OF MALIC ENZYME AND CITRATE CLEAVAGE ENZYME IN LIVER OF THE GROWING CHICK. = =
YEH Y-Y, LEVEILLE GA, WILEY JH.
(LAB. OF NUTR. BIOCHEM., UNIV. OF ILLINOIS, URBANA, ILL.)

---

1. ENZYMES DIET LIPIDS CHICK
2. CHICK DIET LIPIDS ENZYMES
3. LIPIDS DIET ENZYMES CHICK
4. DIET LIPIDS ENZYMES CHICK

In the two references shown in Figure 31, six phrases have been added as subject augmentation from the abstract to the first citation while four have been added to the second reference. In the first citation, for example, let's assume that "embryo" was a truncated word valued at +20. This citation should have added 20 to the other term weights only once to result in the total reference weight and be retrieved. It would, however, have added a term weight of 140 (20 for each appearance in the text term lists as well as the single title appearance). This false count for a single term makes the scheme practically irrelevant and fundamentally unworkable. If a negative weight has been applied, this cummulation could undoubtedly have caused the loss of this citation from the user's output. We discouraged the use of question weighting in CA's service unless it is seemingly required for a specific problem. We feel that the proper use of parameters can for the most part, meet the same requirements more effectively.

The most successful use of term weighting in CA Condensates resulted from the need for negative weighting rather than "not" logic. We would recommend that the phrase concept only be used where search phrase requirements are obvious as in "air pollution" and that only the first word be scanned if the same phrase is to be repeated in inverted form. Any term appearing in the titles itself should never be repeated as a text word.

CODEN

When properly searched, disregarding punctuation, the coden search has proved highly successful and very useful in areas not readily apparent. We always recommend that journal coden be searched in mode 2 but it appears to make no difference in patent or book searches. Journal coden searches have been used for eliminating, through "not" logic, journals which are read regularly by the user and, secondly, to print a contents record within a subject framework for particular foreign journals to which the user's library does not subscribe. Patent and book searches have been also used successfully.

SEARCH DATA ELEMENTS

The availability of additional search capabilities in CA Condensates would be most helpful. Search by molecular formula and registry numbers are available in CBAC (Chemical-Biological Activities), another tape service of Chemical Abstracting Service, but it is not available in any other of their tape services. We feel that all searchable data elements should be included in the CA Condensates tapes since that should be the one all-inclusive CAS tape for retrospective searching and permanent storage. We do not feel that this would result in any lessening of financial resources for Chemical Abstracting Service due to subscription loss any more than that caused by the preparation of

tape services, CA and CT, which also duplicate coverage but for different audiences. We also feel that biological searches would be enhanced by the addition of enzyme commission numbers in all abstracts as keywords.

PRINTOUT FORMAT

The CAS printout format proved highly impractical and was quickly revised by our programmer. For both CT and CA we developed a duplication of identical records on a two column page. This provided the capability for the user to separate the pages and (a) share them with a laboratory assistant, (b) file one copy for future reference while carrying the second copy to the library for research, etc. The double format has been extremely well received. In addition, other relevant changes were made in the two printouts. (Compare Figures 32 and 33 with Figures 27 and 28).

Figure 32 illustrates the changes in the Chemical Titles printout. We have more carefully identified all print capabilities in addition to supplying the header information. Our user, who normally has submitted more than one profile to more than one service tape, can now tell at a glance which service tape is represented and can identify the question and question number from the printout in hand. Spacing also enchanced the readability.

Our CA print format underwent two significant changes before we developed the one shown on Figure 33 which incidentally, is a printout for the prepared profile of Figure 18. This typed copy shows only the right hand column of the double column printout, but does illustrate how well this format lends itself to 8 1/2 x 11 reproduction. To the CAS printout format we have added text keywords which are absolutely necessary for the evaluation of search results. Our CA print format should undergo another adjustment which, with the installation of the new CAS Standard Distribution Format, will be better spaced and allow for quicker observation of each data element.

EVALUATION

After the initial interest stimulated by the seminar has died down, the successful promotion of service and acceptance by prospective patrons often depends on personal contact. We did not have the staff to follow up these seminars properly. As previously mentioned, the profile preparation seminars were given directly to faculty, students, and librarians on eight of the nine campuses, one UCLA Library School group, and to two CIS sponsored groups at UCLA attended by librarians from the five Southern California UC campuses. Only the CIS sponsored group received any actual on-the-spot training in evaluation. On the Riverside campus we were able to give a good deal of individual help, but user feedback cards clearly indicated problems on some other campuses.

CASHION,L QUESTION NO 060 PAGE 1
THRESHOLD WEIGHT 000009 DATE 03/09/70

STABILITY CONSTANTS OF MAGNESIUM CHELATORS

CHEM TITLES

VOL 70 NO 05 REF WT 000014

CCJDAO 1970 70 144
ALKENES FROM EP OXIDES BY REDUCTIVE ELIMINATION WITH MAGNESIUM BROMIDE - MAGNESIUM AMALGAM.=
BERTINI F GRASSELLI P ZUBIANI G

CHEM TITLES

VOL 70 NO 05 REF WT 000014

INOCAJ 0009 70 325-32
PREPARATION OF MAGNESIUM ALUMINUM HYDRIDE. REACTIONS OF LITHIUM AND SODIUM ALUMINUM HYDRIDES WITH MAGNESIUM HALIDES IN ETHER SOLVENTS.=
ASHBY EC SCHWARTZ RD JAMES BD

CHEM TITLES

VOL 70 NO 05 REF WT 000014

JPCHAX 0074 70 281-8
THERMAL DECOMPOSITION OF MAGNESIUM PER CHLORATE AND OF AMMONIUM PER CHLORATE AND MAGNESIUM PER CHLORATE MIXTURES.=
ACHESON RJ JACOBS PWM

CHEM TITLES

VOL 70 NO 05 REF WT 000014

NUPABL 0141 70 257-72
DEPENDENCE OF THE SPECTROSCOPIC FACTORS ON PARAMETERS OF DISTORTING POTENTIALS FOR THE REACTION MAGNESIUM-26(TRITON,ALPHA PARTICLE)MAGNESIUM-25.=
NURZYNSKI J

CASHION,L QUESTION NO 060 PAGE 1
THRESHOLD WEIGHT 000009 DATE 03/09/70

STABILITY CONSTANTS OF MAGNESIUM CHELATORS

CHEM TITLES

VOL 70 NO 05 REF WT 000014

CCJDAO 1970 70 144
ALKENES FROM EP OXIDES BY REDUCTIVE ELIMINATION WITH MAGNESIUM BROMIDE - MAGNESIUM AMALGAM.=
BERTINI F GRASSELLI P ZUBIANI G

CHEM TITLES

VOL 70 NO 05 REF WT 000014

INOCAJ 0009 70 325-32
PREPARATION OF MAGNESIUM ALUMINUM HYDRIDE. REACTIONS OF LITHIUM AND SODIUM ALUMINUM HYDRIDES WITH MAGNESIUM HALIDES IN ETHER SOLVENTS.=
ASHBY EC SCHWARTZ RD JAMES BD

CHEM TITLES

VOL 70 NO 05 REF WT 000014

JPCHAX 0074 70 281-8
THERMAL DECOMPOSITION OF MAGNESIUM PER CHLORATE AND OF AMMONIUM PER CHLORATE AND MAGNESIUM PER CHLORATE MIXTURES.=
ACHESON RJ JACOBS PWM

CHEM TITLES

VOL 70 NO 05 REF WT 000014

NUPABL 0141 70 257-72
DEPENDENCE OF THE SPECTROSCOPIC FACTORS ON PARAMETERS OF DISTORTING POTENTIALS FOR THE REACTION MAGNESIUM-26(TRITON,ALPHA PARTICLE)MAGNESIUM-25.=
NURZYNSKI J

FIGURE 32

CT PRINTOUT FORMAT (UCR)

FIGURE 33

CA CONDENSATES PRINTOUT FORMAT (UCR) PAGE 1

BYRNE/VERITY QUESTION NO 022 DATE 02/18/71

THRESHOLD WEIGHT

COMBUSTION OF METALS

CA CONDENSATES

DIGEST NO 029081R VOL. 74 NUMBER 07 REF WT. 000000
BRXXAA1208692 (CL. A 24D), 14 OCT. 1970, US APPL. 16 JAN 1967 6 PP . .
BRIT.
ALUMINIZED CIGARET WRAPPERS FOR MORE COMPLETE TOBACCO COMBUSTION. ==
LIPPMAN, ALFRED, JR. ARMBRUST, BERNARD F., JR.
LIPPMAN AJR, ARMBRUST, BF JR.
(REYNOLDS METALS CO.)
LIPPMAN, ALFRED, JR. ARMBRUST, BERNARD F., JR.
(REYNOLDS METALS CO.)
COMBUSTION TEMP TOBACCO CIGARET
ALUMINUM CIGARET PAPER BURNING
PAPER CIGARET BURNING ALUMINUM
TEMP COMBUSTION TOBACCO CIGARET
CIGARET TOBACCO COMBUSTION TEMP
TOBACCO COMBUSTION TEMP CIGARET

CA CONDENSATES

DIGEST NO 031001H VOL. 74 NUMBER 07 REF WT. 000000
JAFCAU 0018 0006 70 1161 0005
J. AGR. FOOD CHEM.
DETERMINATION OF BAY 93820 (O-METHYL PHOSPHORAMIDOTHIATE O-ESTER WITH ISOPROPYL SALICYLATE) RESIDUES IN PLANT AND ANIMAL TISSUES BY ALKALI FLAME GAS CHROMATOGRAPHY. ==
THORNTON JS, STANLEY CW.
(CHEMAGRO CORP., KANSAS CITY, MO.).
GAS CHROMATOG ALKALI FLAME
ALKALI FLAME GAS CHROMATOG
BAY 93820 DETN

FIGURE 33 (Continued)

CA CONDENSATES

DIGEST NO 031211B VOL. 74 NUMBER 07 REF WT. 000000
NDPBA6 0022 0006 70 0086 0009
NACHRICHTENBL. DEUT. PFLANZENSCHUTZDIENSTES (BRUNSWICK)
LETTUCE TIPBURN, WITH SPECIAL REFERENCE TO THE UPTAKE OF NUTRIENTS. = =
LEH HO.
(INST. NICHTPARASITAFRE PFLANZENKR., BERLIN-DAHLEM, GER.)

PHOSPHORUS FERTILIZER LETTUCE TIPBURN CALCIUM
CALCIUM LETTUCE TIPBURN FERTILIZERS
NITROGEN FERTILIZER LETTUCE TIPBURN CALCIUM
LETTUCE TIPBURN FERTILIZERS CALCIUM
POTASSIUM FERTILIZER LETTUCE TIPBURN CALCIUM

In October 1970 we hired a biochemically-oriented Ph.D. candidate on an hourly basis to help with some of the particularly difficult profile problems. His normal operating procedure consisted primarily of analyzing the profile with the printout of a current run and developing a procedure and recommendations to be followed in personal consultations. The length and number of these appointments varied with each profile. Local campus users were scheduled for appointments ranging from one-half hour sessions to several one-hour consultations; the latter occasioned by very difficult and exacting literature needs, the unfamiliarity of the terminology or the user's lack of understanding of a computerized term-match system.

Users from other UC campuses were consulted in various ways. A few made appointments and traveled to UCR. A few from UCLA were consulted upon appointment at their location during a quarter break when our subject specialist could go to that campus by intercampus bus. We held one faculty group evaluation session at UC Irvine, also by prior set appointment. Users from other campuses were contacted by the intercampus telephone tie-line in which case brief information or explanation was exchanged (either preceded or followed by written correspondence). This method has been very successful. We hope to continue this approach, perhaps by broadening the scope to include "subject specialists" at the master's level or above on each of the nine campuses.

In general, private explanations of the application of text search strategy were found to be quite helpful. The effectiveness of a search is greatly increased when additional explanation and term expansion is undertaken to include symonyms, related terms, variant spellings, and some abbreviations. The Search Guide provided by Chemical Abstracting Service is particularly useful in explaining the needs for subject authority information and serves as a guide to generic and grammatical variations as well as a check for coverage of broad and narrow terms possibly strengthening the profile's capabilities. This guide, however, appears to be more useful to the information-retrieval specialist than to the subject specialist. It has proven exceedingly helpful in checking left end truncation and as a guide to proper questioning. Personal contact seems to dispel fear of an unknown medium in which the user is conerned with failure. It also serves to remove the worry of excessive costs which will prove too great for him to maintain with present budgetary restrictions. Personal contact has been used to trigger the latent desire for the service, promoted by the seminars or word of mouth, but which has not been acted upon due to other priorities.

Personal contact also points up search effectiveness and shows how it could save the time of the busy researcher while keeping him more fully informed. Personal explanation of how the service acts as a current awareness tool to avoid obsolescence to keep more current in his literature search, and to provide leads to new ideas for operational

improvement makes the activity seem more realistic to the potential user. Additional explanation of the use of the verified citation for speedier Interlibrary Loan service and for the continuance of his literature coverage while he attends professional meetings, takes sabbatical leaves, or basks in the sun on an island vacation is usually quite impressive when weighed against other daily chores. We expect user interest to increase in direct proportion to the extent and amount of personal contact.

EPILOGUE

Some of the problems to date have resulted from resistance within the organization both by users through fear of failure and expense and by the administration in their inability to provide adequate staff and funding. In spite of these odds, great strides have been made. We have a working system for both CA and CT tape services, some experience and expertise in profiling and teaching methods, and a growing number of satisfied users. In addition, we have undertaken research and development on still another tape service BA Previews and BIORI, the tape service names for Biological Abstracts and Bio-Research Index respectively. Salary R&D this year amounted to $9,212.00 and over 2,100 man-hours resulting in an operational program for these services being offered to all nine campuses starting in January 1971.

These new services will be described in detail at a later date after more experience has been gained. A copy of our Search Profile Form (Figure 34) and profile printout (Figure 35) will accompany each printout run made using BA Previews and BIORI tapes supplied through subscription from Biosciences Information Service. Figure 36 illustrates our BA Previews printout format, which does include all header information, and the BIORI printout format is shown on Figure 37. Both of these printout include "type" code symbols which permits quick identification of the information on each line of output. The cost accounting record (Figure 38) is quite similar to the one we use for CA Condensates and CT to provide the user with constantly updated charges until the billing is completed at which time the record is updated and ready to begin again.

We believe we have written an extremely acceptable computer program for the biological tapes and full credit must go to two very bright young men. Mr. Greg Scragg, presently pursuing his Ph.D. in Computer Science at UC San Diego, and Mr. Steve Farrell, an undergraduate student in Computer Science at UC Riverside.

I must mention that my successor in these projects is Mr. Everett Wallace, a librarian who has had wide and varied computer experience covering more than fifteen years. He will have a monumental task for there is still considerable research, development and experimentation necessary on all of the tape services offered at UC Riverside.

FIGURE 34. BA PROFILE FORM - UCR
BIOLOGICAL ABSTRACTS & BIORESEARCH INDEX

SEARCH PROFILE FORM

REQUESTOR: _____

DEPARTMENT: _____

_____

_____

Return to:
Diane Bennett
General Library
University of California
Riverside, California 92507

HEADER CARD			QUESTION:
		QUEST. WEIGHT	
1 4	5	6 9	
	/		

	TERM CARDS						
Question No. (leave blank)	PARAM NO	LOGIC	TYPE	(SIGN)	WEIGH	SEARCH TERM (Keypunch Parenthesis)	
	5 6	7	8	9	12	13 79	80
							DO
							NOT
							USE

FIGURE 35

BA PROFILE PRINTOUT - UCR

QUESTION NO:
NAME:

QUEST. WEIGHT 0

BIOMEMBRANES AND POLYSACCHARIDES

PROFILE PRINTOUT

PARAM	LOGIC	TYPE	WEIGHT	TERM
01	0	T	0	(GOLGI)
01	0	T	0	(DICTOYOSOM
01	0	T	0	(PLASMA LEMMA
01	0	T	0	(PLASMA MEMBRAN
01	0	T	0	(VESICLE
01	0	T	0	(MEMBRAN
02	0	T	0	SACCHARID
02	0	T	0	GLUCAN
02	0	T	0	XYLAN
02	0	T	0	ARABINAN
02	0	T	0	ARABAN
02	0	T	0	GALACTAN
02	0	T	0	MANNAN
02	0	T	0	GALACTURON
02	0	T	0	GLUCURON
02	0	T	0	PECTIN
02	0	T	0	(PECTIC SUBSTANC

FIGURE 36

BA PRINTOUT FORMAT - UCR

QUESTION NO: PAGE 0002
NAME QUEST. WEIGHT 0 10/06/70

CITATIONS

BA VOL. 51 (17) ABSTRACT NO. 088227 WT. 0000
C BICHA/BIOCHEMISTRY
8 (12). 1969 5136-5148
A IRIAS J J/OLMSTED M R/UTTER M E/
T PYRUVATE CARBOXYLASE REVERSIBLE INACTIVATION BY COLD CHICKEN LIVER MITOCHONDRIA
X 02506-10060 10062 10802-10806*14004-23006-
I 10806 85536

BA VOL. 51 ABSTRACT NO. 090894 WT. 0000
C JOBAA/J BACTERIOL
102 (2). 1970 341-346.
A HARTMAN R E/
T CARBON DI OXIDE FIXATION BY EXTRACTS OF STREPTOCOCCUS-FAECALIS-VAR-LIQUEFACIENS PYRUVATE CARBOXYLASE
X 06504 10060 10062 10063 10064-10069 10504 10802-10808-13012-31000*32000
I 31000 07200

BA VOL. 51 ABSTRACT NO. 091943 WT. 0000
C CJBOA/CAN J BOT
48 (4). 1970 777-786.
A BERRY J A/DOWNTON W J S/TREGUNNA E B/
T THE PHOTOSYNTHETIC CARBON METABOLISM OF ZEA-MAYS-M AND GOMPHRENA-GLOBOSA-D THE LOCATION OF THE CARBON DI OXIDE FIXATION AND THE CARBOXYL TRANSFER REACTIONS LEAF MESOPHYLL BUNDLE SHEAT CELLS PLASMODESMATA ASPARTATE MALATE ENOL PYRUVATE STARCH 3 PHOSPHO GLYCERIC-ACID RIBULOSE 1 5 DI PHOSPHATE CARBOXYLASE AUTO RADIOGRAPHY PHOSPHORYLATION GAS EXCHANGE TRANSLOCATION BETA CARBOXYLATION
X 06504 10012 10064 10068 10504 10510 10806 11108 13002-51000 51506*51516 51518 51519-51520 52504 53010
I 51506 25305 25555

FIGURE 37

BIORI PRINTOUT FORMAT - UCR

QUESTION NO:
NAME: QUEST. WEIGHT 0 09/28/70

CITATIONS

BIORI VOL. 70 ABSTRACT NO. 067304 WT. 0000
C ARHEA/ARTHRITIS RHEUM
12 (3). 1969 275-276
A ALARCON-SEGOVIA D/FISHBEIN E/
T ANTIBODIES TO HYDRAZIDE ALTERED SOLUBLE NUCLEO PROTEIN IN TUBERCULOUS PATIENTS RECEIVING ISONIAZIDE ABSTRACT
X 10060 10062-10064 22002-22504-34504-34508*36002-38504-
I 34508 06200 86215

BIORI VOL. 70 ABSTRACT NO. 061832 WT. 0000
C JOMIA/J MICROSC (PARIS)
7 (4). 1968 36A-37A
A CAUTIER A/
T COMPARISON OF SEVERAL METHODS OF EXPOSITION OF COMPLEX DNA HI STONES ON SECTIONS ABSTRACT PANCREAS VIRUS
X 10062-10300 11108-33506*36006
I 33506 03200 85000

BIORI VOL. 70 ABSTRACT NO. 065890 WT. 0000
C AIPBA/ARCH INT PHYSIOL BIOCHEM
75 (1). 1967 169-170
A HOUSSIER C/
T INTERACTION OF NUCLEIC-ACID AND NUCLEO PROTEIN WITH PRO FLAVIN ABSORPTION SPECTRUM
X 10060-10062-10064-10504 10506*

BIORI VOL. 70 ABSTRACT NO. 061820 WT. 0000
C JOMIA/J MICROSC (PARIS)
7 (4). 1968 30A
A DERIEUX J-C/PETITPREZ A/GUILLAUME J/
T ACTION OF THE HISTONES ON BACTERIA AT O CELSIUS AND 37 CELSIUS ABSTRACT
X 10064 13012-23001-31000*
I 31000 06000

FIGURE 38

BILLING RECORD - UCR

NAME: QUEST. WEIGHT 0 09/28/70

BIOMEMBRANES AND POLYSACCHARIDES

BIORI VOL. 70
NAME
BIOCHEM ,RIVERSIDE
FUND NO. -------

QUESTION NO. 0039
BIOMEMBRANES AND POLYSACCHARIDES

THIS ISSUE
NO. OF CITATIONS CHARGES $

THIS BILLING PERIOD
NO. OF CITATIONS CHARGES $

APPENDIX A

THE USE OF CUSTOMIZED LITERATURE SEARCH SERVICE

A SEMINAR

to provide techniques for preparing profiles
for a computer-based current awareness service
for Chemical Titles and CA Condensates.

PRESENTED BY

Kathryn S. Forrest
Assistant University Librarian
University of California
Riverside, California

The magnetic tapes are prepared and provided by Chemical Abstracting Service on contract. This service is available to all University of California students and staff through:

THE LIBRARY
UNIVERSITY OF CALIFORNIA
RIVERSIDE, CALIFORNIA

SEVEN STEPS FOR CODING SEARCH PROFILES*

1. Complete the questioner identification information at the top of the profile form.

2. Write the question in sentence format on the profile form.

3. Determine the essential concepts.

4. Select appropriate search terms for all concepts, including all necessary synonomous and related terms. (Use the Search Guides prepared by CAS and other available reference sources to accomplish term expansion.)

5. Decide on the truncation strategy and the truncation modes for all terms and enter terms on form.

6. Determine the appropriate logic and weighting factors (if used) for all terms and enter codes.

7. Decide on output option and enter appropriate codes.

*Chemical Abstracts Service. Preparation of Search Profiles, Columbus, Ohio.

SEARCH SYSTEM TERMINOLOGY

For Use With Chemical Abstracts Services Magnetic Tapes*

PROFILE: A search profile is a list of terms which describe the interests of the requestor. The profile sets the specifications under which a search item will be regarded, and hence retrieved, as an answer to the search request.

PARAMETER: A parameter is a subset of terms within a single profile. A parameter may use either a single term or multiple alternative terms to represent a concept in the search profile.

"AND" LOGIC: "And" logic is implied for all parameters of a profile. The implied "And" logic requires that a file item must satisfy all parameters before that item is retrieved.

"OR" LOGIC: "Or" logic is applied to all terms within a parameter when it is desired that the presence of any one of these terms in a file item satisfy the parameter in question.

"NOT" LOGIC: The application of "Not" logic to profile terms will cause search items containing such terms to be rejected, regardless of the degree to which other logic is satisfied. All "Not" logic terms in a given profile must be in a single parameter.

WEIGHTING FACTORS: Weighting factors are positive or negative numerical values applies to search terms. The person formulating the profile should attempt to relate the magnitude of the weighting factor for a term to that term's capacity to retrieve relevant information. Negative weights may be employed to simulate "Not" logic when use of the latter would preclude retrieval of relevant file items.

Retrieved file items are sorted for output in order according to the sum of the weights of the profile terms matched in the item; the item with the highest sum is placed first and the others follow in decreasing order. This ordering tends to place the most relevant items at or near the beginning of the output listing.

SEARCH SYSTEM TERMINOLOGY (Continued)

QUESTION WEIGHT: When weighting factors are used, the sum of the weights of all profile terms matched in a search item is compared with a pre-selected threshold weight called the question weight. If the sum of the weights of the matched profile terms exceeds the question weight, the file item is retrieved. Weighting factors are considered only after all parameters have been satisfied.

TERM TRUNCATION: A technique called "term truncation" is used to facilitate retrieval of items containing word fragments which are common to two or more different forms of a word (e.g., between singular and plural forms, or different tenses of a verb). For example, the truncated term *POLYMER* would retrieve the terms POLYMER, POLYMERS, COPOLYMER, POLYMERIZE, POLYMERIZATION, COPOLYMERIZATION, etc. Terms may be searched in any one of four truncation modes.

*Chemical Abstracts Service. Preparation of Search Profiles, Columbus, Ohio.

MODE (Column 8)

SEARCH MODE	SYMBOLIC REPRESENTATION	EXPLANATION
1	*term	Accepts any combination of character preceding the term.
2	term*	Accepts any combination of character following the term.
3	term	Accepts term only if bounded by blank spaces or non-alphabetic characters.
4	*term*	Accepts the term with any string of characters either preceding or succeeding the term.

*The asterisk denotes that a blank space or any character may occupy the position so marked.

TYPE (Column 9)

Column 9 Code	Term Type
A	Authors
C	Coden
M	Molecular formula
R	CAS Registry Number
T	Title or text words

SEARCH PROFILE FORM

Return to: A-6
DIANE BENNETT
GENERAL LIBRARY, ROOM 204
UNIVERSITY OF CALIFORNIA
RIVERSIDE, CALIFORNIA

REQUESTOR: Fill in Name
DEPARTMENT: Fill in Department and/or Grant
CAMPUS: Fill in Campus
FUND: Fill in Department or Grant Fund

Indicate Service
CT
CA Or Both

HEADER CARD

QUESTION NUMBER		OUTPUT INDICATORS	QUESTION WEIGHT
1 3		8 11	12 17
LEAVE BLANK		1,0,0,0	

QUESTION: Original Statement of Request

DETAIL CARD

DUPLICATE	PARAM NO.	NO. OF PARAM	MODE	TYPE	LOGIC	LOGIC	TERMS	TERM WEIGHT
	4 5	6 7	8	9	10	DO NOT KP	11 DO NOT KEYPUNCH ASTERISK 77	78 80
						Question	Weight - Fill in only if term weight is being	
		DUPLICATE					used.	
						Parameter	- We suggest the use of not more than three	
							parameters per question. Several more precise	
							questions prove more successful.	
						Mode - Term Truncation mode 1-4 (see attachment)		
						Type - Five term types are applicable. Any combination		
							of term types may be used in a single profile.	
							(see attachment)	
						Logic -	"Or" logic is indicated by O, "Not" logic is	
							indicated by N. "And" logic is implied for all	
							parameters and is not coded for individual terms.	
							All "Not" logic terms in a given profile must be	
							in a single parameter.	
						Term Weight	- These may be positive or negative numerical	
							values. Weighting factor is considered only	
							after all parameters have been satisfied. The	
							sum of the weights of the matched profile terms	
							must exceed the question weight to be retrieved.	
						All High	Order Zeros must be coded. Alpha O is written Ø	